Functional failure sequences in traffic accidents

By

Deniz Atalar

A Doctoral Thesis

Submitted in partial fulfilment of the requirements

for the award of

Doctor of Philosophy of Loughborough University

Acknowledgements

The completion of this thesis has come with the help of colleagues and co-workers in the field of road safety. First of all I would like to thank my wife Hatice and my parents, whose great support allowed me the patience to finish this document. I would also like to thank my close friends who provided encouragement and support when necessary.

Special thanks needs to be given to Professor Pete Thomas whose supervision and support during this PhD made this document possible. Also thanks needs to be given to Professor Andrew Morris who provided guidance with regards to the necessary additions to this thesis. I would also like to especially thank my internal examiner Dr Richard Frampton whose constructive feedback contributed greatly to this thesis being completed.

I would like to thank Mr Russell Danton and Ms Claire Quigley who both provided their case coding for the comparison study and took the time and effort to explain to me the different accident causation methods used in this thesis. I would also like to thank Mr Pete Paynton, Mr Duncan Pearce and Mr Robert Gisby for taking me out on site and initiating me in the way of the world of in-depth accident investigation. Furthermore I would like to thank Mr Rob Newton for taking me out to retrospective studies on site.

Thanks needs to be given to Dr Pierre Van Elslande and Mr Christophe Perrin who kindly and patiently explained the methodologies that they developed and use. Special note must be made to Mr Julian Hill for allowing me to use the comparison data in the comparison study and providing me the use of the OTS data as necessary.

Thanks also needs to be given to Mr Jim Lenard who provided a portion of the SPSS files for use in this study and to Dr Alun Owen with whom our discussions of statistical procedures with the use of multivariate datasets greatly benefitted the outcome of this thesis.

This thesis uses accident data from the UK Co-operative Crash Injury Study (CCIS). CCIS was managed by TRL Limited on behalf of the UK Department for Transport (DfT) Transport Technology and Standards Division who funded the project along with Autoliv, Ford Motor Company, Nissan Motor Company and Toyota Motor Europe. Daimler–Chrysler, LAB, Rover Group Ltd, Visteon, Volvo Car Corporation, Daewoo Motor Company Ltd and Honda R&D Europe (UK) Ltd have also funded CCIS. The data were collected by teams from the Birmingham Automotive Safety Centre of the University of Birmingham, the Vehicle Safety Research Centre at Loughborough University, TRL Limited and the Vehicle & Operator Services Agency of the DfT. For further information about the UK collision investigation data please go to RAIDS@dft.gsi.gov.uk

Abstract

This thesis examines the interactions between road users and the factors that contribute to the occurrence of traffic accidents, and discusses the implications of these interactions with regards to driver behaviour and accident prevention measures.

Traffic accident data is collected on a macroscopic level by local police authorities throughout the UK. This data provides a description of accident related factors on a macroscopic level which does not allow for a complete understanding of the interaction between the various road users or the influence of errors made by active road users. Traffic accident data collected on a microscopic level analysis of real world accident data, explaining why and how an accident occurred, can further contribute to a data driven approach to provide safety measures. This data allows for a better understanding of the interaction of factors for all road users within an accident that is not possible with other data collection methods.

In the first part of the thesis, a literature review presents relevant research in traffic accident analysis and accident causation research, afterwards three accident causation models used to understand behaviour and factors leading to traffic accidents are introduced. A comparison study of these accident causation coding models that classify road user error was carried out to determine a model that would be best suited to code the accident data according to the thesis aims.

Latent class cluster analyses were made of two separate datasets, the UK On the Spot (OTS) in-depth accident investigation study and the STATS19 national accident database. A comparison between microscopic (in-depth) accident data and macroscopic (national) accident data was carried out. This analysis allowed for the interactions between all relevant factors for the road users involved in the accident to be grouped into specific accident segmentations based on the cluster analysis results.

First, all of the cases that were collected by the OTS team between the years 2000 to 2003 were analysed. Results suggested that for single vehicle accidents males and females typically made failures related to detection and

execution issues, whereas male road users made diagnosis failures with speed as a particularly important factor. In terms of the multiple vehicle accidents the interactions between the first two road users and the subsequent accident sequence were demonstrated.

A cluster analysis of all two vehicle accidents in Great Britain in the year 2005 and recorded within the STATS19 accident database was carried out as a comparison to the multiple vehicle accident OTS data. This analysis demonstrated the necessity of in-depth accident causation data in interpreting accident scenarios, as the resulting accident clusters did not provide significant differences between the groups to usefully segment the crash population. Relevant human factors were not coded for these cases and the level of detail in the accident cases did not allow for a discussion of countermeasure implications.

An analysis of 428 Powered Two Wheeler accidents that were collected by the OTS team between the years 2000 to 2010 was carried out. Results identified 7 specific scenarios, the main types of which identified two particular ‘looked but did not see’ accidents and two types of single vehicle PTW accidents. In cases where the PTW lost control, diagnosis failures were more common, for road users other than the PTW rider, detection issues were of particular relevance. In these cases the interaction between all relevant road users was interpreted in relation to one another.

The subsequent study analysed 248 Pedestrian accidents that were collected by the OTS team between the years 2000 to 2010. Results identified scenarios related to pedestrians as being in a hurry and making detection errors, impairment due to alcohol, and young children playing in the roadside. For accidents that were initiated by the other road user’s behaviour pedestrians were either struck after an accident had already occurred or due to the manoeuvre that a road user was making, older pedestrians were over-represented in this accident type.

This thesis concludes by discussing how (1) microscopic in-depth accident data is needed to understand accident mechanisms, (2) a data mining approach using latent class clustering can benefit the understanding of failure

mechanisms, (3) accident causation analysis is necessary to understand the types of failures that road users make and (4) accident scenario development helps quantify accidents and allows for accident countermeasure implication discussion. The original contribution to knowledge is the demonstration that when relevant data is available there is a possibility to understand the interactions that are occurring between road users before the crash, that is not possible otherwise. This contribution has been demonstrated by highlighting how latent class cluster analysis combined with accident causation data allows for relevant interactions between road users to be observed. Finally implications for this work and future considerations are outlined.

Keywords: **accident causation, driver perception, multivariate analysis methods, latent class clustering**

Table of contents

List of tables

List of figures

Abbreviations

ACASS Accident causation analysis with seven steps

AIC Akaike information criterion

BIC Bayesian information criterion

DfT Department for Transport

CIREN The Crash Injury Research and Engineering Network

DREAM Driving reliability and error analysis method

FARS Fatal Accident Reporting System

GIDAS German In-depth Accident Study

GOF Goodness of Fit

HFF Human Functional Failure method

ITS Intelligent Transport Systems

KSI Killed and Seriously Injured

LOC Loss of Control accident

LCC Latent Class Cluster analysis

NASS National Automotive Sampling System

PTW Powered Two Wheelers

QIE Quasi Induced Exposure method

R R statistics package

RTA Road Traffic Accidents

SPSS Statistical Package for Social Scientists

TRACE Traffic Accident Causation in Europe

TRRL Transport and Research Road Laboratory

UK OTS UK On the Spot Accident Study

VRU Vulnerable Road Users

WHO World Health Organization

Glossary

UK OTS The UK's On the Spot (OTS) accident data collection project was an in-depth accident research project that aimed to collect accident data on the scene minutes after they were reported as occurring by the police, and gather all relevant perishable data with regards to the accident. This project was carried out between the years 2000 to 2010.

STATS19 The STATS19 data is the national data source on traffic accidents in Great Britain. It provides detailed information on vehicles involved in accidents that are resulting in injuries and can be used for different research aims within these contexts. The data is collected yearly by different police local authorities and provided by the DfT to researchers.

SafetyNet was a sixth framework European Union funded project aimed at the development of a new European Road Safety Observatory to help in the development of safety policies by providing data and knowledge. This project was carried out between the years 2006 to 2008.

TRACE was a sixth framework European Union funded project aimed at accident causation analysis and the evaluation of safety benefits of technologies in terms of traffic safety. This project was carried out between the years 2006 to 2008.

1 Introduction

The purpose of this chapter is to clarify the theme of this thesis and prepare the reader for the main ideas that follow. An introduction to some of the important themes and explanation of the aims and objectives of this thesis will be made in this chapter.

This thesis will look at identifying a novel approach to understand accident data using a statistical methodology that will help both further understand the nature of a traffic accident compared to past methods and help researchers by identifying particular types of traffic accident scenarios within the United Kingdom (UK), that are more prevalent and problematic than other accident types. This comparison will be demonstrated by developing accident scenarios through in-depth and national statistical data analysis for the identification of possible traffic safety countermeasures. The introduction and literature review that follows aims to develop the research questions that this thesis broaches and provide a sound basis to identify the topics that will be tackled in the analysis chapters and discussion/conclusion.

1.1 Traffic accidents as a global problem

A road accident is defined by United Kingdom (UK) law as any occurrence on the public highway (including public footway) where at least one vehicle collides with another vehicle, another road user, or a stationary roadside object (DfT, 2006). A traffic accident is a complex phenomenon which, to some level, involves the road user, vehicle, environment (road structures) and other individuals that are within the environment (drivers, pedestrians, cyclists or riders). Though most elements within the environment are constant and unchanging the presence of road users constantly interacting with other road users and the road environment creates a continuously fluctuating dynamic system. Accidents are usually a result of a combination of the above involved participants and system interacting and creating an accident situation (Allnutt, 1987).

In the year 2004 over 1.2 million people died worldwide as a result of road traffic collisions, an average of 3,242 fatalities per day, and an estimated 20 million to 50 million people were injured or disabled. Traffic accidents were estimated to have cost UK £334 billion annually around the world (WHO, 2009).

The World Health Organization (WHO) estimates that road traffic injuries were the 11th leading cause of death worldwide, accounting for 2.1% of all deaths globally and for 23% of all injury deaths in 2006 (WHO, 2006). Figure 1 shows that road traffic injuries are projected to become the 5th leading cause of death by 2030 as mobility increases in emerging economies and long standing diseases and other health impacts are mitigated (WHO, 2006) (World Health Organization, 2009). In 2007, 91% of fatalities occurred in low-income and middle-income countries despite records showing that these countries contain only 48% of the worlds registered vehicles (WHO, 2009).

The first report on road casualties in Great Britain carried out by the Department for Transport (DfT) recorded I4.7 million licensed vehicles and 178,000 injuries accidents in the year 1951 (DfT, 2006). In 2005 32.9 million vehicle and 268,000 injury accidents were reported and in 2007 182,000 traffic accidents occurred in the UK, 3,307 of these resulting in fatalities, 28,000 resulting in serious injuries and 217,000 in slight injuries (DfT, 2008). Despite the number of vehicles increasing by 700%, the number of accidents increased by 20% between these years (DfT, 2008).

Of the accidents occurring in the UK in 2013 the three Vulnerable Road User (VRU) groups (pedestrians, pedal cyclists and motorcyclists) between them accounted for almost 50% of all deaths and 60% of all seriously injured casualties (DfT, 2013). This is similar to the figures reported by the WHO, VRU groups made up 46% of all global fatalities in road safety throughout the world in 2004 (WHO, 2009).

According to UK figures motorcycle users, are roughly 35 times more likely to be killed and over 50 times more likely to be seriously injured per mile ridden, in a road traffic accident than car occupants. Pedestrians and pedal cyclists are roughly 11 times more likely to be killed and cyclists are 30 times more likely to be seriously injured per mile walked and cycled respectively in a road accident than car occupants (DfT, 2013). Improvements in terms of safety for these groups are important in order to significantly drop Killed and Seriously Injured (KSI) casualty figures.

TOTAL 2004

RANK	LEADING CAUSE	%
1	Ischaemic heart disease	12.2
2	Cerebrovascular disease	9.7
3	Lower respiratory infections	7.0
4	Chronic obstructive pulmonary disease	5.1
5	Diarrhoeal diseases	3.6
6	HIV/AIDS	3.5
7	Tuberculosis	2.5
8	Trachea, bronchus, lung cancers	2.3
9	Road traffic injuries	2.2
10	Prematurity and low birth weight	2.0
11	Neonatal infections and other	1.9
12	Diabetes mellitus	1.9
13	Malaria	1.7
14	Hypertensive heart disease	1.7
15	Birth asphyxia and birth trauma	1.5
16	Self-inflicted injuries	1.4
17	Stomach cancer	1.4
18	Cirrhosis of the liver	1.3
19	Nephritis and nephrosis	1.3
20	Colon and rectum cancers	1.1

TOTAL 2030

RANK	LEADING CAUSE	%
1	Ischaemic heart disease	12.2
2	Cerebrovascular disease	9.7
3	Chronic obstructive pulmonary disease	7.0
4	Lower respiratory infections	5.1
5	Road traffic injuries	3.6
6	Trachea, bronchus, lung cancers	3.5
7	Diabetes mellitus	2.5
8	Hypertensive heart disease	2.3
9	Stomach cancer	2.2
10	HIV/AIDS	2.0
11	Nephritis and nephrosis	1.9
12	Self-inflicted injuries	1.9
13	Liver cancer	1.7
14	Colon and rectum cancer	1.7
15	Oesophagus cancer	1.5
16	Violence	1.4
17	Alzheimer and other dementias	1.4
18	Cirrhosis of the liver	1.3
19	Breast cancer	1.3
20	Tuberculosis	1.1

Figure 1: Leading causes of death, 2004 and 2030 (estimated) (adapted from WHO, 2009)

1.2 The road user in the traffic environment

The safe and successful operation of the roadway system is influenced by a number of user factors, including (1) physical or physiological factors, (2) psychological or behavioural factors and (3) cognitive factors. The road user's limitations in terms of experience, impairment, physical and mental skills, motivation and other characteristics are factors in the safe and efficient functioning of the roadway system (Olson, 2002).

Risk awareness is fundamental for safe driving. A driver can plan ahead if they are aware of present and potential risks to eliminate or reduce them (Lidestam, Lundqvist, & Rönnberg, 2010). The visual sense is the main channel of sense in operation during a driving behaviour for the road user, and is used in order to correctly identify the traffic situation, make a decision and perform the required behaviour (Riemersma, 1979). Drivers focus on objects in their visual field by routinely moving their eyes in order for that object to come onto the centre of the retina (Herslund & Jørgensen, 2003). In complex scenarios this may prove impossible and relevant information may not be perceived leading to an incorrect manoeuvre or behaviour and in turn to a potential crash situation.

Large scale accident studies have focused on identifying human error in the accident environment since the 1960's onwards (Carsten, Tight, Southwell, & Plows, 1989; Ljung, 2007; Sabey & Staughton, 1975; SafetyNet, 2009; Treat et al., 1979; Van Elslande & Fouquet, 2007); Historically human errors were found to be the main factor in 70-80% of accidents in in-depth accident causation studies (Sabey & Staughton, 1975; Treat et al., 1979). Human causes are historically viewed as the triggering variable that influences accidents. They are interpreted with regards to the prominence of specific risk factors within the traffic environment.

Human error research currently takes a systematic approach to observe the errors that road users make to consider the latent conditions that contribute to these errors within the system (Delhomme, Dobbeleer, Forward, & Simões, 2009). The nature of these errors and conditions are tied to the attributes within the traffic environment. Reason (1990) contributed that human error is inevitable thus it is necessary to understand the steps that lead to the errors occurrence, this can only be done by understanding the latent factors present during the behaviour and drawing conclusions from the relationship between these factors.

1.3 Accident causation research

The aim of accident causation research is to identify the source of accidents and ultimately reduce or eliminate them (Lehto & Salvendy, 1991). This type of research is used to understand how an accident happened by placing the features in a causal chain link creating a timeline from pre-event to the post-event, by analysing the accident data after the event has happened. In traffic accident research accident causation analysis aims to understand the interactions that occurred between the human road user and the other elements within the traffic environment in order to understand the main failure sequence that occurred.

Accident causation analysis requires the analyst to relate causal relationships by using an accident causation model to infer the different failures and contributory factors that are present for a road user during the build-up to and occurrence of an accident. A main failure is the main reason that an accident occurred, whereas a contributory factor can be defined as all factors that contributed to an accident occurring. This analysis requires the accident investigator to interpret the data and group the different factors that are both present and also inferred by the analyst. This requires a level of subjectivity to be included in the analysis despite the best wishes of the analyst.

Azencott et al. (2007) reviewed the use of causality in accident research and how statistical relationships can be understood underneath this umbrella. The subjective nature of the coding makes the step from statistical association to causal relationship a question of interpretation. Criteria need to be found in order to distinguish causal from non-causal associations. The method used to gather data is particularly important as the causal relationships between the factors depends on its validity and applicability.

Azencott et al. (2007) identified two essential points in the interpretation of causation data:

1. There should be a strong correlation/association between the possible cause and the possible effect.

2. The causation must precede the impact. In terms of accident research: There should be clear evidence that exposure to the risk factor preceded the accident. Another important aspect is that there should be a plausible explanation.

One of the main aspects of analysing accident data is interpreting interaction (between people variables). Due to the large number of variables present within the data and the different types of data fields it is necessary to understand how the factors interact and interpret the analysis within the scope of how multiple factors are interacting.

It is also important that the relationships presented are plausible. As causation is an explanation of a process as relationships are uncovered and deduced from the data. As the nature of causation data is subjective it is necessary to take these features into consideration and make the data collection process as objective and uniform as possible.

1.4 Real world accident data

Historically, different types of study methodologies have been used to understand the behaviour of the road-user in an accident – more specifically how road user behaviour may have contributed to the accident.

The different approaches include;

1. In-depth accident investigations conducted at the scene of the accident (within a few minutes of crash occurrence)
2. In-depth accident investigations conducted retrospectively (within a few days of the accident occurrence)
3. Studies using self-report follow-up techniques with those involved
4. Observation studies/Naturalistic driving studies

In-depth accident investigations conducted at the scene of an accident aim at obtaining all relevant data at the scene of an accident, these studies provide information that is otherwise not available in other studies (Hill & Cuerden, 2005). In-depth on scene accident investigation studies send a team of crash investigators to the crash site as soon as possible to examine and collect volatile data from the site, as well as information regarding the environment and interview data from the individuals involved within the crash. Although the data acquisition process has been considered excessive considering the research findings (Grayson & Hakkert, 1987), the amount of data obtained is quite large and suitable for a number of different analysis possibilities.

When an in-depth accident investigation occurs evidence is gathered in three ways, either factual evidence is obtained on the spot, interview based evidence is obtained from the road users, or an assessment is made by the road safety professional on the spot (Sabey & Staughton, 1975). Interview based data is the least reliable of the three methods of data acquisition as individuals are not necessarily inclined to give truthful information in regards to the accident occurring though the interviewer is able to make observations on the road users reliability due to their experience in interviewing individuals in these types of situations (Sabey & Staughton, 1975). Road users that participated in the accident can reprocess memories and come up with a different interpretation of the accident which would put themselves in the position of the non-contributory road user to the crashes occurrence.

The advantage of in-depth accident data is that the level of detail collected is relatively high compared to other methods and the factors can be related to the crash outcome in detail. In-depth data is usually collected by independent research teams consisting of expert investigators that use a strict methodology collecting key variables with regards to the accident, human road user (including interviews), vehicles involved, injury data, road infrastructure and scene environment information, accident reconstructions and accident causation analysis, this data is in turn analysed (Hagstroem et al., 2010). The disadvantages for this data is the high cost of obtaining the data, possible issues in sampling and representivity of the data, and the relatively small size of the data collected in many cases.

The data collection process by the in-depth collection activities is aimed to be independent, compared to national accident data collection procedures which are influenced by judicial investigation aims and procedures, although impartial in nature these investigations will be carried out with the judicial system in mind and will aim at assigning blame to certain vehicles. In-depth accident data collection aims at improving safety, not who was to blame (Hagstroem et al., 2010). This data can be used in a multitude of areas such as vehicle design and crashworthiness, policy and legislation, child occupant safety and road infrastructure (Hagstroem et al., 2010).

In-depth accident investigations conducted retrospectively help supplement already available reports through analysing the accident and/or vehicles to have a more detailed view of the accident (Langweider, Hummel, & Mueller, 1997). In-depth retrospective studies researchers go to recovery yards and take data from vehicles that have been involved in a crash usually a few days after the event. They then use this data as well as police reports to reconstruct the accident. This data can be used to understand driver error as long as the police data contains relevant information related to the drivers behaviour as well as trip data where possible, though retrospective data is found to develop less statistically significant findings compared to on the spot data (Ranney, 1994).

Self-reports are methods that are designed to gather self-reported information from road users in terms of their traffic behaviours. They can include methods, including questionnaires and inventories, interviews, focus groups and driving diaries (Lajunen & Özkan, 2011; Tivesten, Jonsson, Jakobsson, & Norin, 2012). Questionnaire studies aim at gathering retrospective information from the road users that were involved in the accident. Accident data from insurance reports provide self-reported information from the individuals that are involved in the crash. This data is a representation of the accident according to the individual’s perception. In both surveys and insurance data the recollection of the accident is particularly important as this is the base of the information that is provided. According to Lajunen & Özkan (2011) basic motor and perceptual processes are difficult to measure with self-report methods since the driver is unaware of most of the automated processes that are carried out while driving.

Social desirability is another source of bias in self-reported data (af Wåhlberg, 2010; Lajunen & Özkan, 2011), which can be described as "a tendency to give answers that make the respondent look good" (Lajunen & Özkan, 2011, p. 55). A distinction can be made between impression management (lying) and self-deception. Impression management tend to increase in public compared to anonymous settings, while self-deception is more linked to personality (Lajunen & Özkan, 2011). Interview data, survey data and insurance data is likely to be influenced by the perception and retrieval of the original event by the road user.

Observation studies aim to document the frequency or occurrence of behaviour while driving, rather than to understand the direct cause of behaviour (Eby, 2011). These types of studies are particularly useful when researchers are aiming to observe a behaviour that is directly observable such as seat belt use in a population. Naturalistic driving studies are forms of experimental studies that collect data from instrumented vehicles by using video recordings and computers to acquire information about road user's behaviours while using their vehicles. A number of data sources can be used such as simple accelerometers as well as more advanced systems such as eye tracking devices (Dozza, Bärgman, & Lee, 2013).

This can be done by taking videos of the road user for each specific trip and/or by collecting data about the physical behaviours that the road user is making and how the vehicle is reacting. These studies have been used to collect data on the safety aspects of different in-vehicle systems or to better understand the driver behaviours that result in crashes, near misses or conflict situations in the road system. The driver is placed in real-world conditions and instrumented vehicle studies are conducted to assess driver behaviour under natural conditions where the driver is facing normal traffic conditions on their normal routes (Klauer, Perez, & McClafferty, 2011). Though this data is very rich in nature the amount of crashes collected are relatively small. In Dingus et al.'s (2006) 100-Car Study continuous data were collected on 109 vehicles for a minimum of 12 months and the resulting data set included 69 crashes, 761 near-crashes, and 8295 incidents. This data source is rich in providing multiple levels of information with regards to the driver's behaviour during driving but in terms of accident data the cost is relatively large for a small number of accident cases.

National accident statistics provide information on what type of vehicles are involved in different types of accidents and other crash characteristics. This data is usually obtained from police reports and put on a national database (Langweider et al., 1997). This information is used to monitor national progress from year to year and to understand the different type of accidents that occur on national roadways. The detail level is below that of in-depth retrospective and on scene in-depth accident studies.

Table 1 demonstrates the type of information collected by different types of data collection methods and the cost of each case. In-depth on scene accident studies are quite expensive to conduct though the level of data collected is higher compared to the other methodologies. Retrospective in-depth accidents are less costly than in-depth on scene studies but still require a large amount of man power to collect all relevant data. Studies using self-report data have quite a low cost but have limitations in terms of validity.

Table 1: Traffic accident studies data analysis types and data

Analysis type	On Scene	Self-report	Interview relating to accident	Volatile data	Cost of data per case
On scene in-depth accident investigations	Yes	No	Yes where possible	Yes	High
Retrospective In-depth accident investigations	No	No	Retrospective if possible	No	High
Studies using self-report data	No	Yes	No	No	Low
Observation studies/ Naturalistic driving studies.	Yes	No	No	If possible	Low to Medium/ Medium to High

Observation studies also can have a lower cost compared to other study types and provide more valid data than self-report studies, while naturalistic driving studies have a higher cost and provide very detailed data for a smaller sample size.

1.5 Methods used to analyse accident causation data on driver behaviour

When analysing a large number of accident cases it is possible to draw inferences using different statistical methods. There are three types of procedures used in data analysis, these are (1) frequency distributions to understand the numerical values for the collection, (2) using exploratory data analysis methods to get to know the data and (3) inferential statistics to understand what relationships and conclusions can be made from the data (Howell, 2009).

When analysing traffic accident causation data traditionally, each individual case is collected and analysed coding the qualitative information and contributory factors according to a classification scheme (Sandin, 2009). A case study examining accident records provides insight into the traffic safety situation using information from accident cases but is considered more in line with an intuitive rather than a data-driven approach (Kweon, 2011). These results are then analysed using frequency analysis and correlations. A large number of accident causation studies have combined individual accident cases into aggregate cases to identify particular types of accident scenarios (Sandin, 2009; Ljung Aust, 2010; Treat et al., 1979, Sabey & Staughton, 1975).

A scenario can be named as a prototype or model of an accident process characterised by a sequence of events that contribute to damage either to individuals or their property/environment (see, for example (Cicioni, Giuliano, Castellano, & Lattanzi, 1994; di Marzo, Almenas, & Gopalnarayanan, 1995; Karwat, 1992)). This studying of the past experiences can be used to develop prevention strategies to chains of events (Fleury & Brenac, 2001).

A typical scenario of an accident is defined by Van Elslande and Fouquet (2007) as “the typical progress with which we can connect a group of accidents which present resemblances from the point of view of the chain of the phenomena, whether they are analysed from an historical, a functional or a causal point of view” (p. 4). These constructions can be considered as a prototype, as they are attained from a number of similar accidents rather than only one accident (Fleury & Brenac, 2001).

If we consider the process of an accidents occurrence as the presence of factors that are different from the norm, then the prevention of these behaviours will stop an accidents occurrence. When taking the factors into consideration it is necessary to attribute the factors to not just the accident that has happened but to the traffic environment and infrastructure as a whole. The introduction of countermeasures to broach these effects can then contribute to traffic safety benefits.

1.6 Countermeasures

A road safety measure is a device, program or tool whose main purpose is to improve road safety (Elvik, Hoye, Vaa, & Sorensen, 2004). The theoretical approach that is used to collect and analyse road safety data determines the countermeasures that will be provided and applied. We can divide the use of countermeasures according to the three main proponents within the traffic environment.

1. The road user
2. The vehicle
3. The environment/infrastructure

The countermeasures that have been used to address these proponents have historically counted on the three E’s. According to Damon (1958) the then director of the Kansas City Safety Council gave a presentation where he presented a drawing of a triangle with sides labelled education, enforcement and engineering (as cited in Groeger, 2011). Through the years a number of

E’s have been added to this model, most notably emergency medical response.

The road user is the interactee in the traffic system that suffers the consequences of any conflicts in said system. Countermeasures aimed at road users focus on education, enforcement and measures to reduce accident risk.

Elvik et al. (2004) identified 124 countermeasures according to studies that listed their effectiveness. The studies that were included either provided numerical estimates or stated the number of accidents that the study was based upon.

With regards to vehicle countermeasures the main methods of precautionary measures during the 70’s, 80’s and 90’s were concentrated on secondary safety and passive safety measures (airbags, seatbelts) that are aimed at reducing or preventing injury after a crash has occurred (Evans, 1991). The active safety approach has been incorporated during the latter 90’s and 00’s and in present day integrated safety measures are aimed to be incorporated in vehicles. The ‘active safety’ approach is traditionally associated with technologies that are likely to result in crash avoidance (e.g. Intelligent Speed Adaptation (ISA), Enhanced Stability Programmes (ESP) and Lane Departure Warnings (LDW)) (Morris et al., 2006).

The 21st century has seen the development of advanced driver assistance systems (ADAS) that aim to improve comfort, safety and convenience by either assisting the driver or by taking over certain driving tasks (Richardson, Barber, King, Hoare, & Cooper, 1997; Young, Birrell, & Stanton, 2011). These systems are aimed at influencing vehicle control and either monitoring the vehicle performance and providing assistance, or monitoring the driver’s performance and providing a warning of impending danger/dangerous driving. The system works by providing a reactionary driving behaviour either fully or by adding to the driver’s behaviour. However, these kinds of automation may actually cause an increase in reaction time and situation awareness (Brookhuis, de Waard, & Jannsen, 2001). According to Lee (2004) the limitations, preferences and conditions of the driver needs to be taken into account for the collision warning system to be successful. In the future with further development of a multitude of systems these issues will be arising furthermore and the manner in which these systems interact with the driver will be of critical value. Another issue of note is how to train older drivers in the use of these systems.

At the fundamental level the two critical traffic elements that can be influenced by roadside objects are the traffic (vehicle) speed and lateral positioning (De Ridder et al., 2006). Engineering countermeasures aim at providing improvements in terms of the vehicle (build, performance, occupant protection, passive safety) and improvements to the environment through providing clues to the road user that signal appropriate driving speeds and lane positioning for the type of road and indicate expected behaviours within the traffic environment (for other road users)

Safety technologies that have been developed to help road users with regards to crash avoidance and/or mitigation currently fall under four possible headings (Atalar et al., 2012). These technologies aim at tackling the issues discussed above:

1. Passive safety measures: These measures reduce the consequences of an accident by managing the crash forces. Passive safety refers to the vehicles protective measures when involved in an accident (e.g. seat belts, airbags).
2. Active safety measures: These measures reduce the possibility of accidents occurring by taking preventative measures. Active safety normally involves the implementation of safety technology within the vehicle which is specifically designed to reduce the risk of a crash occurring (e.g. vehicle braking and stability, electronics).
3. Integrated safety measures: These measures aim at integrating active and passive safety systems within a vehicle to allow the vehicle to adapt to a pre-crash situation and either stop the crash from occurring or reducing the crash consequences by reacting to the crash appropriately.

4. Rescue safety measures: also known as tertiary technologies. These measures optimise the rescue phase by supplying information on crash severity and location to rescue services.

1.7 Summary

In this chapter a brief description of traffic accidents as a global problem was identified, with an overview of the current and future projected rate of traffic accidents in terms of injury and monetary loss. The role of the road user in this environment and a brief identification of how human failure occurs were carried out. The collection procedures of accident causation data was explained as a data source for understanding driver behaviour.

The different ways of collecting accident data on scene and retrospectively (in-depth on scene accident data, on scene national accident data, retrospective accident data, insurance data, questionnaire studies and observation studies/naturalistic driving studies) and their uses were explained and compared.

A review of the problems in relation to road traffic accidents has been carried out and discussed with a brief description of countermeasures with regards to traffic accident data and the different types of safety technologies available.

2 Literature review

2.1 Introduction

The aim of this thesis is to identify the most prominent human failures that occur in a traffic accident by identifying the human, vehicular and environmental/infrastructural factors that contribute to an accident, in order to develop a better understanding of how traffic accidents occur. Previous work has concentrated on identifying risk factors that have been linked to increased accident risk. This thesis aims to take a holistic approach, to identify the interactions between risk factors and focuses on the interaction between driver failure and other factors present in the traffic system. This process will also investigate the links between different risk factors and identify these links in different accident scenarios. Human error has been historically quoted as the main accident cause of up to 90% of accidents (Treat et al., 1979; Sabey & Staughton 1975), but a more thorough understanding of how these errors occur concurrently with other factors is not currently available. In order to understand the interactive nature of traffic accidents an interactive ergonomics model was used within this thesis, guided by a human factors perspective. This approach helped develop a unique way to analyse risk factors that are present during traffic accidents.

The review of the literature is based on the following assumptions:

- A detailed understanding of human behaviour and failure in traffic accidents can contribute to a better understanding of accident causation.
- There are four main possible types of factors that can cause or contribute to a traffic accident occurring. These factors are human factors, vehicular factors, environmental factors and infrastructure factors.
- Accident causation methods are the most systematic way of understanding human failure within traffic accidents.

- A need to use statistical analysis methods with in-depth and national data is necessary to form a better understanding of accident sequences

2.2 Acquisition of driving skills

In order to understand how a road user interacts with the road environment a detailed understanding of both the acquisition of the driving skills as well as a model to understand how driving behaviours occur is beneficial. The acquisitions of skills related to driving are acquired over time. Fitts and Posner (1967, as cited in Evans 1991, 101) identified three stages during the acquisition of driving skills (learned through trial and error); An early (cognitive) stage, an intermediate (associative) phase and a final (autonomous) stage. These stages can be viewed in figure 2. When transferred to a perspective considering driving in the early stage the road user learns to understand the components that are necessary for the driving behaviour to take place, similar to a way a toddler learns to walk. The individual is acquiring connections within the brain to form a cohesive map of individual components to the driving task (Evans, 1991).

In the intermediate stage strategies are employed so that the individual is aware of outputs from other drivers and individuals within the traffic environment, though still giving a high level of attention to the driving task (Evans, 1991). In the autonomous phase the attention level to detail is lessened and a lot of the reactions are now done without extra thought being given to the action (Evans, 1991). The driver accumulates knowledge which in turn leads to drivers forming expectations of how to react to specific driving situations, and as these expectations increase visual search patterns and behaviours become automatic. Progress through these three stages or changes in the knowledge on which performance directly occurs as a result of practice, virtually all of the skills that have been systematically studied show a gradual slowing of improvement as task experience increases.

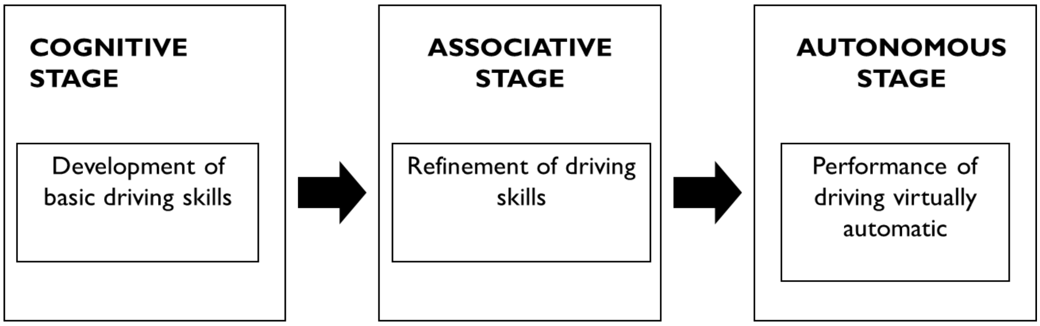


Figure 2: Fitts and Posner's three stage model (Adapted from Fitts & Posner, 1967)

Research carried out in a driving simulator has shown that novice drivers scanning behaviours are less complex than experienced drivers (Underwood, Chapman, Bowden, & Crundall, 2002), and they are also less able to remember events other than central driving events (Underwood, Chapman, Berger, & Crundall, 2003). However when they received training in terms of visual search patterns their performance gradually increased though not to the level of more experienced drivers (Chapman, Underwood, & Roberts, 2002). Traffic behaviours need to be learned gradually, as drivers experience different driving situations their experience and skill will increase which in turn lessens the amount of unconscious errors that they make (Bjørnskau & Elvik, 1992; Underwood et al., 2003).

The rate of improvement will slow faster for easy tasks than for hard tasks (Groeger, 2002). This in turn allows drivers to concentrate on the aspects of the specific driving situation that they are confronted with in terms of what is most relevant (Koustanaï, Boloix, Van Elslande, & Bastien, 2008). As the driving environment is extremely complex, drivers can only gather a limited amount of information from what is occurring around them. Each driver will select what is significant to them depending on knowledge, the situation and trip objectives (Van Elslande & Fouquet, 2007).

While road users move along a route they continuously perceive the environment and in turn relate these movements in a general setting. The individual's route direction depends on the understanding that individuals develop as they move along this path. The level of detail with which they describe their surroundings will be determined by how well they have perceived their environment (P. E. Michon & Denis, 2001). According to Michon and Denis (2001), the three most important reasons why landmarks are required during navigation are: 1) signalling where an action should be executed, 2) creating the link to the next section of the route, and 3) reassuring navigators that they are still on track. This understanding of the navigation of what they have experienced allows road users to map these reasons onto any network graph that represent route navigators. The objective of a person generating route directions is to deliver these descriptions in such a way as to allow a high level of understanding of the environment described. The nature of information acquisition and learning in the traffic environment follows a similar pattern.

In order for the driving behaviour to be carried out appropriately a driver needs to select the most relevant information from the traffic environment, in order to make the necessary driving actions. After selecting the necessary information, then an interpretation of the information needs to be made in order for a decision process to be carried out. These decisions are based on previous knowledge of different situations the road user has been confronted with (Van Elslande & Fouquet, 2007).

2.3 Models related to driving

Historically in road safety a number of models have been proposed to explain the process of how a human failure occurs. The two main ways of identifying driving behaviours can be grouped into descriptive models and functional models. The descriptive models describe what the driver does while the functional model describes the motivational and situational factors that lead to the driver making their decision (Oppenheim & Shinar, 2011).

Descriptive models are analytical rather than predictive, and aim to describe what the driver has to do either as a part of what they do or as a whole

(Carsten, 2007). They can be divided into hierarchical models and control loop models (Oppenheim & Shinar, 2011).

The hierarchical task model of Michon is one of the main types of descriptive models (Carsten, 2007). Driving as a task is a skilled activity with several hierarchies (Summala, 1996). The generalized problem-solving task of the user can be further divided in three levels of skills and control: strategical (planning), tactical (manoeuvring), and operational (control) respectively (J. A. Michon, 1985).

The driver can also engage in other activities besides the driving task (e.g., talking on the phone, daydreaming) that can be described according to the three task levels as well. Risk is related to in terms of driver's non-performance of a manoeuvre leading to a problem occurring rather than in terms of the quality of the driving task (Carsten, 2007).

We can describe the driving task according to control loop models by using inputs, outputs, and feedback (Oppenheim & Shinar, 2011). These models of driving have traditionally been expressed either in terms of guidance and control or in terms of human factors. There is difficulty in using these models to understanding driver behaviour with regards to complex behaviours with modern cars (Oppenheim & Shinar, 2011).

Functional models focus on the cognitive state of the driver and psychological functions to help understand it (Oppenheim & Shinar, 2011). Informational processing models include the driver as a passive information channel that performs different acts within capacity limitations. The driver is shown to go through different stages based on perception, decision and response selection (Wickens, 1992). Two of the crucial components within this model are attention allocation mechanisms and the feedback loop (Oppenheim & Shinar, 2011). Individual differences in relation to behaviours have been shown to affect future performance parameters. The feedback loop implies that data processing is a continuous process and as new stimulus is entered new modifications and processes are made (Shinar, 2007).

Motivational models of driving aim to describe how the driver manages risk or task difficulty (Carsten, 2007). Theories that determine risk in regards to behaviours undertaken are most commonly used in this model. Drivers are assumed to take the amount of risk that they are willing to endure for each behaviour (Gibson, 1966; Oppenheim & Shinar, 2011).

When interpreting these behaviours accident rates and measures of injuries are used, though it is necessary to understand this process in terms of drivers' behaviours and performance measures (Oppenheim and Shinar, 2011). Driver behaviour is usually measured by response time. Response time is typically composed of at least three components: (1) perception reaction time – how long the driver needs to perceive input and decide on a response, (2) movement time – how long the physical movement takes, and (3) how long the device requires to carry out its response (Oppenheim & Shinar, 2011). If we consider an individual pressing a brake, they first will perceive that a situation arises that requires the brake pedal to be pressed, they then motion to press the brake pedal and once they have performed this movement the machinery will move to cause the vehicle to brake.

2.4 Crash sequences and accidents

2.4.1 Theoretical approaches to accident investigation (Historical)

When analysing accidents investigation practices have historically relied on a number of different methodological and practical perspectives to help determine and distinguish accident causes, and tie them together with possible countermeasure considerations. Accident investigation practices usually provide a model to frame the way that an accident happened and how they can be prevented. Benner jr. (1985) found differences in the performance of 17 evaluated US accident investigation methodologies. Benner jr. (1985) stated that accident models should be realistic, definitive, satisfying, comprehensive, disciplining, consistent, direct, functional, non-casual, and visible.

This analysis type mirrors the types of traffic accident investigation that has been carried out throughout the years. Lundberg, Rollenhagen & Hollnagel (2009) grouped previous accident investigation into four possible areas based on investigation types through the years and broadening complexity. These models are;

1. Simple linear system models (cause–effect models): These models approximated on preventing the most obvious cause. Looking for actions that seemed incorrect and correcting those.
2. Complex linear system models (epidemiological models): These models highlighted factors that may not be observable immediately and contributed to the occurrence of said accident. Best known today as the Swiss Cheese model (Reason, 1997).
3. Complex interactions: Which Reason (1997) described as the discrepant casual chains where managerial activities at the ‘‘blunt end” could lead to latent conditions at the ‘‘sharp end”.
4. Performance variability: When the environment and the system itself changes in a system performance the variability of performance is required and may lead to negative effects. To avoid these it is necessary to concentrate on a new equilibrium for the system (Sundström & Hollnagel, 2006).

(Elvik et al., 2004) identified five different theories that have been historically used to try to explain accidents, particularly with reference to road safety analysis. They are;

1. The theory of accidents as purely random events
2. Statistical accident theory and accident proneness theory
3. Causal accident theory
4. Systems theory and epidemiological accident theory
5. Behavioural accident theory

The timeline in relation to these theories prominence and use can be viewed in figure 3. The theory of accidents as random events aimed to explain why there is variation within a group of individuals when accidents should be completely random (Elvik et al., 2004). The human participant was viewed as taking part in the accident on a completely random basis. The likelihood of an accidents occurrence was viewed as not being related to the behaviours of the participant.

The view of accidents as random events was shaken when Greenwood and Yule (1920) discovered that certain workers were responsible for most of the accidents in munitions factories. As this could not be due to randomness they put forward the idea that certain people were more prone to accidents compared to others. This theory was predominant between the years 1920-1950 and pertained that certain individuals contributed to most of the accidents that occurred (Elvik et al., 2004).

Causal accident theory aims to find the real causes of accidents which in turn it identified as being multi-causal events with multiple factors leading to the accident rather than having one single cause (Elvik et al., 2004). This model uses a holistic approach to integrate all possible factors and events leading up to an accidents occurrence.

Systems theory aims to modify the technical components of the road transport system in order to match the road environment requirements to human capabilities (Elvik et al, 2004). A system can be considered in the context as a group of activities that are constantly interacting. Any type of breakdown in this interaction will lead to an accident occurring.

Behavioural accident theory in turn concludes that human risk assessment and acceptance is what determines how many accidents happen and this can only be altered by changing the target risk (Elvik et al., 2004).

Human assessment and risk assessment determines the number of accidents that each society has. The only way of lowering this number is by having a target of a lower number.

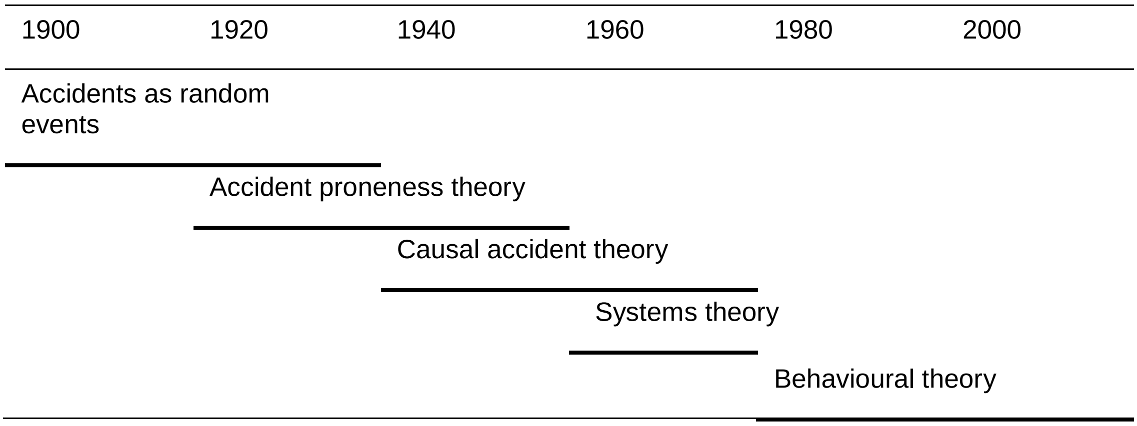


Figure 3: The heyday periods of various accident theories (Adapted from Elvik et al., 2004)

In concordance with the above models road safety management has become progressively more ambitious over time, and in the results desired. Breen (2012) identified four phases in road safety management as evident since the 1950's:

- Phase 1: During the 1950's it was assumed that direct educational and training approaches could more or less solve the problem. As the WHO states, these measures provide general support, but there is little or no evidence to indicate casualty reduction effects for this approach. It is notable how easy it seems to be to slip back into this emphasis through political expediency, industrial demand or lack of professional challenge.
- Phase 2:1960's-1970's focused on system-wide interventions guided by the 'Haddon Matrix'. William Haddon Jr. developed a model based on a public health perspective model, dividing a multitude of topics into factor areas (personal factors, vector or agent factors, physical and environmental factors and social environmental factors) and categorizes groupings (pre-event, event, and post event) to explain different phases of a typical crash situation. Using this methodology it is easier to determine during which crash phase a factor relating to a crash occurs, and to in turn take necessary precautions. This

framework for road safety aimed for intervention on infrastructure, vehicles and users in the pre-crash, crash and post-crash stages but did not yet bring in institutional management (Breen,2012). Table 2 demonstrates a typical Haddon matrix model.

Table 2: Haddon Matrix

Crash phase	Human	Vehicle	Environment
Pre-Crash	Physiological factors, Psychological factors	Active safety systems	Road conditions, Traffic laws, Environmental conditions
Crash	Physical stature, Seat belt use	Impact type, Passive safety systems	Roadside characteristics
Post-Crash	Physical stature, Medical condition	Passive safety systems	Rescue safety performance

- Phase 3:1980's-1990's focused on system-wide interventions, targeted results and institutional leadership. Lead agencies in good practice countries used action plans with headline targets to be achieved with evidence-based packages of measures.
- Phase 4: From the mid-1990's onwards focused on system-wide interventions; long-term elimination of serious health loss, supported by interim targets, shared responsibility and strengthened institutional delivery. This is the perspectives of 'zero vision' or 'durable safety' (Rumar & Wegman, 1999) which underline the notion of responsibility shared between the road users, the society that builds and maintains roads, as well as the industry that conceives and sells the means of transportation.

These interventions have renewed emphasis on speed management; better road, and vehicle crash protection; the present day theories concentrate on using the Safe System approach and behavioural models as to identify and rectify issues in traffic safety.

The Safe System approach for the management of road safety have evolved over the last few decades in developed countries. This approach recognises that mistakes and errors will be made by the human road user in the transport system. The Safe System approach aims to provide a road system design that accounts for these human errors to stop any serious or fatal injuries to the road user from occurring. A Safe System approach (OECD, 2008) has the following characteristics:

- It recognises that prevention efforts notwithstanding, road users will remain fallible and crashes will occur.
- It stresses that those involved in the design of the road transport system need to accept and share responsibility for the safety of the system, and those that use the system need to accept responsibility for complying with the rules and constraints of the system.
- It aligns safety management decisions with broader transport and planning decisions that meet wider economic, human and environmental goals.
- It shapes interventions to meet the long term goal, rather than relying on “traditional” interventions to set the limits of any long term targets.

A number of countries have adopted Safe System measures, or measures that are similar to a Safe System approach. Vision Zero based in Sweden, The Netherlands Sustainable Safety Strategy, The Australian Safe Systems strategy and the UN decade of action for road safety all used similar vernacular.

Salmons, McClure and Stanton (2012) suggested that rather than using a completely systems theory based approach the methods adopted need to be shifted towards a more detailed approach to the transport system. They further suggest that human factors based approaches to accident analysis provide the basis for a detailed system based approach that is necessary.

2.5 Accident causation research

The aim of accident causation research is to identify the source of accidents and ultimately reduce or eliminate them (Lehto & Salvendy, 1991). This type of research is used to understand how an accident happened by placing the features in a causal chain link creating a timeline from pre-event to the post-event, by analysing the accident data after the event has happened. It is used in a diverse number of disciplines with closed environment settings, such as construction work and nuclear plants, and allows for investigators to understand the specific factors that lead to an accident happening. This understanding is aimed to lead to a formulation of a plan or way to stop the accident from happening in the same manner in the future.

As the traffic environment is a large open environment it is difficult to directly transfer the same methodology used within closed environments without making fundamental changes to the analysis methods. Most research in terms of traffic accidents looks at factors in a very detailed manner by identifying the types of effects these factors have in terms of an accident or specific traffic manoeuvres occurring.

As traffic accidents are partly random events there is difficulty analysing traffic accidents before they happen or as they happen, so in-depth accident studies are used to analyse crashes after they occur, using physical data acquired after the crash as well as interviews and questionnaire data to reproduce events leading to the crash.

In order to understand crashes it is necessary to put them into the context of a theory. Accident causation models aim to provide a theoretical framework to identify the source of accidents and ultimately reduce or eliminate them (Lehto & Salvendy, 1991). Figure 4 demonstrates an analysis of accidents that leads to solutions and the variables that are included in this analysis according to both a top down and a bottom up procedure. This type of analysis takes the road users' needs into account and uses empirical study based data to estimate and evaluate the effectiveness of different countermeasures used.

The two general approaches that have been used in the study of causation in traffic crashes are the expert/clinical method and the statistical method (Blower & Craft, 2005). The clinical method relies on the expert to determine the cause of a particular crash, involving a team of multidisciplinary experts studying crashes using the members expertise in all relevant disciplines of road safety to analyse the primary and contributing causes for each specific case, the data of which can then be examined using case review and statistical methods (Blower & Craft, 2005).

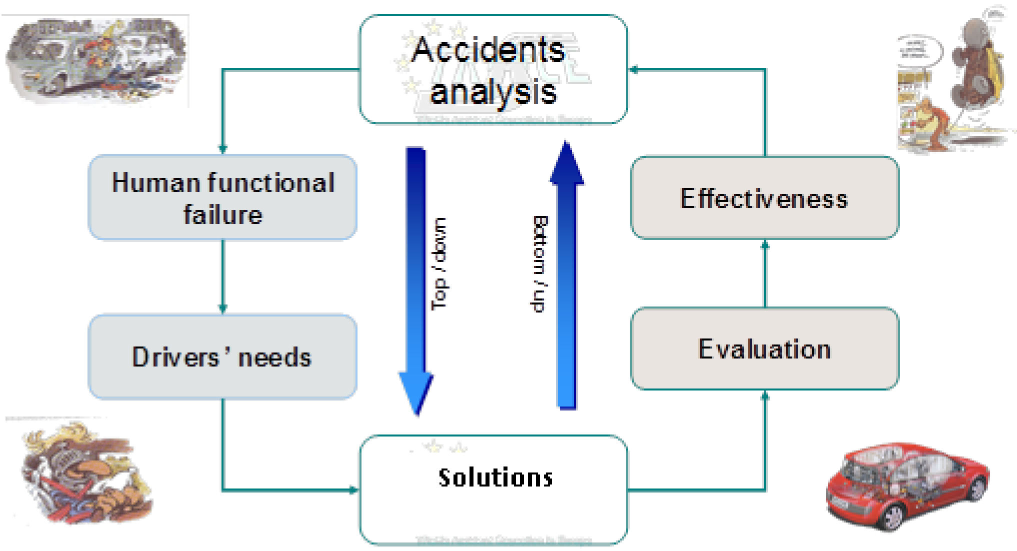


Figure 4: Accident analysis and solution situations (Adapted from Hermitte, 2012)

The statistical method defines the cause of a crash not by assigning it but by demonstrating how it changes with regards to risk. In this approach a cause is a factor that increases risk (Blower & Craft, 2005). Expected risk in the statistical method can be measured in either absolute or relative terms. When exposure measures are available the absolute risk of a crash can be calculated. For example, if travel estimates for vehicles and trucks are available, the absolute rates of crashes can be calculated, and the crash risks per mile travelled for the two different vehicles can be compared. When exposure information is not available, conditional or relative risk is calculated using the acquired crash data (Blower & Craft, 2005). When aiming to identify countermeasure solutions, it is necessary to use a systems perspective as the basis of identification.

2.6 Human error

Human error can be identified in a number of different ways. The most common ways of identifying human error is to identify the types of errors that road users can make. Three main perspectives are currently prevalent with regards to understanding human error in traffic behaviours: Norman's (1981) error categorization; Reason's (1990) slips, lapses, mistakes, and violations classification; and Rasmussen's (1982) skill, rule, and knowledge error classification (Oppenheim & Shinar, 2011). Norman (1981) broke down slips into three major categories; (a) error in the formation of the intention (falsely classifying the situation or not specifying the situation), (b) faulty activation of schemas (actions intruding when not expected), and (c) faulty triggering (by blending the components of actions or false triggering of behaviours).

Reason's (1990) 'Swiss Cheese' model differentiates between active failures and latent failures during transport accidents. Any failures related to deficiencies in the management system are referred to as latent failures (e.g. poor road design). Latent failures require psychological precursors to be present to turn into active failures. Unsafe acts or active failures errors are identified as slips, mistakes and violations. Slips were identified as behaviours that road users make without meaning to, such as taking a turn automatically towards your house even though your path is different, mistake are behaviours that you did not mean to make, such as taking a wrong turn and violations are behaviours that are meant to break a safety rule on purpose, such as running a red light or speeding.

Rasmussen (1982) identified human error as either man-machine or man-task misfits, system or frequent misfits being system errors and occasional misfits either related to man or the system being either human error or system failures. Rasmussen (1982) further identified three levels of knowledge for the human behaviour as rule based, knowledge based or skill based.

Rasmussen defined these behaviours with regards to their complexity and regarded them as being hierarchical. The first level are those related to the skills that the road user needs to learn in order to put this into action, certain skills are more important for driving than others. These behaviours are smooth and automated and do not take place with conscious control (Rasmussen, 1982). The second level is related to the rule based complexity, it is controlled by the middle level of the hierarchy. The behaviour that is carried out is based on goal oriented behaviour and requires understanding and analysis of the environment and behaviour (Ramussen, 1983). The third level is knowledge based and is the highest level of hierarchy and requires a mental model for this process to work. The individual develops a useful plan for an action to carry out the behaviour, particularly when faced with unfamiliar situations (Rasmussen, 1983). Rasmussen (1982) determined that as technical systems are designed for very definite reasons, the ultimate aim of the human in man-made systems is as important as causal explanations based on engineering analysis.

2.7 High risk factors

Accident precursors that are in place within the traffic environment, contribute to an accident by making it possible for them to occur. When considering accident causation, systems approach and behavioural approaches, all of these models have factors interacting with one another with regards to the triggering and cause of the accident. The assumption of risk in the traffic environment is one that is commonly studied in road traffic safety studies. Some of the main factors that are considered to affect drivers adversely are fatigue, drugs, age, gender and in car distractions (Hole, 2007).

When identifying the human road user within the traffic environment it is necessary to make a distinction between the different senses and possible errors that can be attributed to a change in these senses or a failure to comprehend these senses. Drivers need to be aware of what is taking place in the road and the surrounding traffic (Merat, Jamson, Lai, & Carsten, 2012). Situational awareness is defined by Endsley (1995) as the operator (driver) being aware at an advanced level of the situational understanding and projecting future system states. Though identifying danger depends on the situation that a driver is confronted with, if a driver is confronted with a single vehicle on a deserted road this is easier to process and react to than a group of vehicles on a busy road (Koustanaï et al., 2008).

The expectation of a driver also influences visual search patterns, for example when a road user has right of way they pay less attention to vehicles that are coming from the direction that does not have right of way and this, at least theoretically, brings a possibility that a crash can occur from that direction. Certain driving behaviours require different expectations and reaction times, for example an overtaking manoeuvre requires the driver to both understand the speed of the vehicle ahead, the distance needed for an overtaking behaviour to occur and also the time before the arrival of a vehicle that is coming in the opposite direction. The same manoeuvre occurring on a motorway though is less complex as there is no need to make the calculations on the approaching vehicle, it is only necessary to identify whether the lane on the right (or left) is free for this manoeuvre to take place.

The visual sense is particularly important during driving and when considering the road user it is necessary to understand the different types of stimulation that can either cause s/he to guide their attention away from the task at hand (lose vigilance). These can be identified as either inattention or distraction.

An important factor that requires a deep level of understanding is distraction of the road user. Regan, Hallett and Gordon (2011) identified the different types of distraction as relating to our senses and also to our thought process; diversion of attention toward things we see, things we hear, things we smell, things we taste, things we feel, and toward things we think about (internal distraction).

Yannis, Laiou, Papantoniou, and Gkartzonikas (2013) further identified the difference between inattention (lack of attention, insufficient attention, cursory attention selection of irrelevant information, orienting of attention on internalised thoughts and daydreams, engagement in activities secondary to driving, symptoms of drowsiness, looking away from the forward roadway) and distraction (diversion of attention away from driving, or safe driving, competing activity, inside or outside the vehicle, driving-related or not the competing activity may compel or induce the driver to divert attention toward it; safe driving is adversely effected).

Stutts et al. (2005) analysed data coded as distraction and inattention in the US Crashworthiness Data System (CDS), and reported that in terms of overall event durations the most common distractions were eating and drinking (including preparations to eat or drink), distractions inside the vehicle (reaching or looking for an object, manipulating vehicle controls, etc.), and distractions outside the vehicle (often unidentified). Distractions decreased driving performance and were found to have a relationship, as measured by longer periods of no hands on the steering wheel, with distraction towards the outside of the vehicle, and not staying in lane. Staubach (2009) identified distraction and reduced activities as an influence on all cases in causing error in an in-depth investigation carried out on 474 cases in Germany using GIDAS data.

Studies show that the driver groups at highest risk of crash involvement are younger (17-25) and older drivers (65 years and older) (Clarke, Ward, Truman, & Bartle, 2007; Clarke, Ward, & Truman, 2005; Lardelli-Claret et al., 2011). Younger drivers tend to have 2.5 times the rate of accidents compared to older drivers when all variables (relative number in the population, amount of drivers on the road) are controlled for (Clarke et al., 2007). Lardelli - Claret et al. (2011) detected that compared to middle age drivers (45-49 years) youngest (18–20 years) and oldest drivers (60–64 years) had a higher crash risk. The main issues with young road users was identified as risk-taking and lack of skill and with older drivers it was perceptual problems and difficulty judging and responding to traffic flow. Using a sample of police reported traffic accidents in the US state of Alabama and the Crash Analysis Reporting Environment (CARE), McGwin Jr. and Brown (1999) identified that with respect to crash characteristics, older drivers were less likely to have crashes involving driver fatigue, during the evening and early morning, on curved roads, during adverse weather, involving a single vehicle, and while traveling at high speeds. Older drivers were over-represented in crashes at intersections and/or involving failure to yield the right of way, unseen objects, and failure to heed stop signs or signals. Crashes occurring while turning and changing lanes were also more common among older drivers. Alcohol was less likely to be a factor in traffic crashes involving older adults.

Speeding as a factor in accidents has been well-researched and attributed to an increase in the injury severity level of road users (Elvik et al., 2004; Taylor, 2000). Using multivariate logistic regression on the US Fatal Accident Reporting System (FARS) data Bédard, Guyatt, Stones, & Hirdes (2002) found that “travelling at a speed of 112 kph (70 miles per hour (mph)) or more was independently associated with an 164% increase in the odds of a fatality compared with speeds of less than 56 kph (35 mph)” (p. 725). Bédard et al. (2002) also pointed out that the larger the deceleration of the vehicle the higher the likelihood of post-injury medical complications, independent of age and injury severity.

Analysing 3,437 UK police accident reports (including 1,296 in detail) Clarke et al. (2005) found that underlying factors in regards to younger driver accidents are risk taking factors rather than skill factors. Despite having good driving skills they take unnecessary risks, this causes them to be confronted with more dangerous situations which in turn leads to more accidents occurring. They also found that 50.4% of young driver’s accidents came in the hours of darkness, accidents involving aggressive driving, driving while over the alcohol limit, and inappropriate or illegal speed all show an increase for young drivers, especially males, during night driving hours (Clarke et al., 2005).

2.8 Previous real world studies to gather real world accident data

In-depth accident studies have been carried out from the 1960's onwards in the UK and throughout the world. In the UK, the first study that included on scene work was carried out by Starks and Miller in 1961 at the Road Research Laboratory (RRL) and continued with Mackay in 1964 who formed a multi-disciplinary team working closely with the Birmingham Accident Hospital. These studies investigated issues related to passive safety, vehicle design and accident causation throughout the 1960's (Morris, Smith, Chambers, & Thomas, 2005).

Transport and Research Road Laboratory (TRRL) studies

This study was carried out between the years 1970 to 1974. A multi-disciplinary team of researchers from the TRRL were on-call 24 hours in and around the Transport and Road Research Laboratory in South East Berkshire, UK and were called on scene by the emergency authorities immediately on receipt of a notification of an incident (Morris et al., 2006). All information with regards to volatile data such as skid marks, debris and position of the vehicles involved after the impact were collected and interviews were conducted (Morris et al., 2006). In total, the team investigated 2,130 road traffic accidents (RTA), which represented 60% of all injury accidents reported to the police in the area. Analysis of the data revealed that the survey was representative of the South East Berkshire area but not of the country as a whole (Morris et al., 2006).

In this study road users were analysed in terms of fault of the accident and were divided into three levels; (1) primarily at fault, (2) partially at fault, and (3) victim. Drivers were found to be at fault in 40% of the accidents, partially at fault in 19% of the accidents and a victim in 39% of the accidents. Pedestrians were found to be primarily at fault in 65% of accidents, partially at fault in 14% and a victim in 21% of the accidents (Sabey & Staughton, 1975).

Human factors were found to contribute to 95% of accidents and were identified as the sole contributor in 65% of accidents. Road environment factors were attributed to be contributory in 569 accidents (28%) and vehicular factors in 8.5% of accidents (Sabey & Staughton, 1975).

Tri level Study

Treat et al. (1979) analysed accidents on three levels looking at police report data, traffic accidents on site and using a multidisciplinary team to analyse cases. A total of 2,258 on site and 420 in-depth accidents were collected between the years 1970 to 1975 in the Monroe area of Indiana, USA. The Tri level Study was not nationally representative as it did not have a statistical design and was conducted in one state specifically.

Treat's studies utilised essentially the same methodology as Sabey & Staughton’s (1975) study with a slight modification such that a three-level approach to data collection was used so that each factor was allocated as being a 'definite', 'probable' or 'possible' factor in the causal chain of events leading to an traffic accident (Morris et al., 2006).

In the Tri level study human factors were definitely causative as the main factor 70.7% when the accident was reviewed in-depth, definite or probable in 92.6% of accidents and 64.3% at the initial on site review of the accident. The environment was definitely causative as the main factor in 12.4% of in-depth accidents and 18.9% of on-site accidents and the vehicle was the causative factor in 4.5% of in-depth accidents and 4.1% of on-site accidents (Treat et al., 1979). These main factors identified for all of the accidents can be seen in Table 3.

Table 3: Tri level study main factors (Adapted from Treat et al., 1979)

Human factors	%	Vehicular factors	%	Environmental factors	%
Looked but did not see	17.6	View obstructions	3.8	Brake systems	2.9
Inattention	9.8	Wet roads	3.8	Tires and wheels	0.5
Excessive speed	7.9	Design problems	1.9	Body/door openings	0.5
Improper Manoeuvre	6.2	Transient hindrances	1.9	Communication systems	0.2
Internal distractions	5.7	Control hindrances	1.2	Steering systems	0.2

GIDAS study

The German In-Depth Accident Study has been collecting data of in-depth investigations on scene in the Hannover area since 1973 and in the Dresden area since 1999 in Germany (Hautzinger, Pastor, Pfeiffer, & Schmidt, 2007). The team consists of doctors and technicians investigating traffic accidents involving injured persons by a statistical spot-check procedure.

Damage to vehicles, accident traces and injuries are documented in detail and the injury classification AIS (Abbreviated Injury Scale) is used to describe the injury severity of each occupant (Morris et al., 2006).

This study calculates Equivalent Energy Speed (EES) and speed change at impact (Delta-V) using vehicle deformation data and reconstructs the kinematic of vehicle and passengers. The maximum avoidance speed of the crash is also calculated. Accident causation data is also interpreted by the accident investigators after the accident using the information and interview data collected on scene.

INRETS Study

The Department of Accident Mechanisms at the Institut National de Recherche sur les Transports et leur Securité (INRETS) conducted an in-depth study of traffic accidents in the Salon de Provence region of France between the years 1980 to 1987 (Morris et al, 2006). Over the course of this study, 400 RTAs were examined in detail (Girard, 1993).

The INRETS team was alerted to an RTA at the same time as the emergency services and collected as much information as possible at the scene of the crash. The study concentrated on ‘vanishing’ data such as skid marks, final rest locations of vehicles involved, weather and roadway conditions. This data together with preliminary assessments of the vehicles were collected by a trained technician, whilst a psychologist interviewed the driver either on the scene or as soon as was possible afterwards. A second phase of the study was undertaken subsequent to the accident comprising an investigation of the demographics of the driver, investigation of details of the journey being undertaken and a technical vehicle investigation (Morris et al., 2006).

The main aim of this study was depth of information collected rather than national representivity, so a small number of cases were collected and analysed in-depth with each case being analysed on scene and afterwards by a group of investigators and decided upon in detail. This analysis has been continued since 2006 in the Salon de Province region. Accident causation data has been collected in this study.

ITS study

The Institute of Transport (ITS) in Leeds analysed 12,554 injury accidents in North Leeds in the year 1988, on urban roads that consisted of speed limits of 40 mph or less (Carsten et al., 1989). This study was carried out by first observing police reports and the researchers following this by administering questionnaires to the respondents either by interview or post. The response rate was 50% to the questionnaires. Almost 70% of the accidents that were observed, were found to occur on junctions. Site visits were also undertaken for this study and case conferences where the case contributory factors was determined by the investigators, was analysed with two investigators and entered into the database. This study used a four step analysis for causation that coded cases in terms of:

1. The immediate failure that precipitated the event
2. A failure that increased the likelihood of the accident happening
3. The road user behaviour or lack of skill that lead to those failures
4. The explanation for the failure or behaviour

ANCIS study

The ANCIS (the Australian National Crash In-depth Study) is a collaborative research program involving the automotive manufacturing industry, State and Federal Government agencies and automobile associations. This study was started in 2000 and collects retrospective data with a focus on injuries and fatalities (Fildes, Logan, Fitzharris, Scully, & Burton, 2007).

This study collects in-depth data on a representative sample of passenger vehicle crashes in Victoria and New South Wales on severe crashes where the injury to at least one occupant results in their being hospitalised. The analysis process is started at the hospital where suitable participants are located and retrospective interviews, site visits and vehicle analysis are carried out. Over 1,000 accident cases have been collected since the start date, and the analysis is particularly concentrated on passive safety system development and human error analysis,

NASS study

The National Automotive Sampling System (NASS) is a national study that uses probability based sampling to collect data in 60 locations around the US. NASS collects 55,000 cases per year and uses a statistical weighting method to represent the 6.2 million annual crashes that are reported to the police. NASS data has collected a sample of 150,000 crashes since 1979 that include minor, moderate, serious, and fatal crashes (NHTSA, 2010a).

NASS has two parts: the Crashworthiness Data System (CDS) and the General Estimates System (GES). These systems work by randomly selecting police accident reports at police agencies. For the CDS researchers gather interview and medical record data to add to the on-site investigations that they have carried out. GES choses approximately 60,000 crashes each year that reflect the geography, roadway mileage, population, and traffic density of the U.S for sampling purposes.

CIREN study

The Crash Injury Research and Engineering Network (CIREN) collect injury causation data by using a multidisciplinary approach to crash data. This data is collected within eight centres throughout the US, and each individual's injury is linked to the crash mechanism that caused it allowing for a deeper understanding of how crashes occur and how to prevent prospective injuries (NHTSA, 2010b).

NMVCCS study

The National Motor Vehicle Crash Causation Survey (NMVCCS) aimed to collect crash causation data to compliment data acquired from vehicles, the roadways, and the environmental conditions. The data was collected on crashes involving light vehicles, during the period January 2005 to December 2007 throughout the US (NHTSA, 2010b). This study collected information at the crash scene and used a two-dimensional sampling frame reflecting on both space and time of crash occurrence in sampling crashes from among those occurring between 6 a.m. and midnight (NHTSA, 2010b).

FICA study

The Factors Influencing the Causation of incidences and Accidents (FICA) study was carried out between the years 2003-2006 in Sweden. This study was led by Chalmers University Vehicle and Traffic Safety Centre (SAFER) as a collaboration partnership between Volvo Car Corporation, Saab, AB Volvo, Trafiverket (the Swedish Transport Administration) and Autoliv. FICA collected approximately 200 cases during this time period but mostly concentrated on collecting single vehicle and intersection accidents (Ljung Aust, 2010; Ljung, 2007). This study used an accident causation system to collect all relevant factors in the crash and collected telephone interview data from individuals involved in the crash.

Table 4 demonstrates the different types of accident data collection carried out in these studies, the data collection procedures that were carried out, if an accident causation system was used in the coding of this data and the sampling design that was undertaken. Reviewing the different studies it can be seen that the studies using accident causation analysis usually collected a lower number of cases, had a regional sampling design and mainly were on scene accident analysis studies.

Table 4: Real world studies by data type, causation method and sampling design

Study	Data type	Accident causation coding system data	Sampling Design
TRRL	On scene	Yes	Regional
Tri level	On scene	Yes	Regional
GIDAS	On scene	Yes	Regional
INRETS	On scene	Yes	None
ITS	Retrospective	Yes	Regional
ANCIS	Retrospective	No	National
NASS	On Scene	No	National
CIREN	Retrospective	No	National
NMVCCS	On scene	Yes	Weighted
FICA	On scene	Yes	Regional

2.9 Case sampling with real world data

Sampling design can be defined as rules that determine the sampling units that are included in the sample (Hagstroem et al., 2010). Case sampling is particularly important when using real world in-depth accident data, as the applicability of countermeasures is determined by how representative of the population the data collected is. Different types of data require different types of sampling protocols.

There are two main types of sampling, probability and nonprobability sampling. Probability sampling is random in some manner and represents the population, while non-probability sampling is used when population representivity is not necessary or possible. There are four major types of probability sample designs:

1. Simple random sampling
2. Stratified sampling
3. Systematic sampling
4. Cluster sampling

Table 5: Comparing sampling designs

Sampling procedure	Sampling type	Pros	Cons
Simple random sampling	Random selection	Exactly random	Difficult to carry out for accident data
Stratified sampling	Samples from separate time sequences (groups)	Low cost, Low error variance	Difficult to identify exact strata
Systematic sampling	Systematically select samples in an ordered sequence	Evenly sampled population	The sample may be compromised in extreme cases
Cluster sampling	Probability sampling using clusters of elements	Easy to use for specific areas rather than whole population	Representivity may be low, High error variance

Table 5 illustrates the different sampling procedures and pros and cons for each of the different sampling procedures. Hagstroem et al. (2010) reviewed the different types of sampling procedures with regards to in-depth accident data procedures and short discussions of these procedures are provided below.

2.9.1 Simple random sampling

In simple random sampling, it is assumed that the sample is chosen completely at random from the population of interest, and that every element within the population has an equal likelihood of being selected within the sample (Aczel, 2009). This form of sampling allows each accident to have an equal chance of being investigated, though this sampling procedure would allow a chance for all accidents to be studied it would be an extremely small chance (Hagstroem et al., 2010).

When considering in-depth accident data the number of accidents that occur during different time and day periods vary year to year so an exact random sample would not be possible without actually accessing all available accidents.

2.9.2 Stratified sampling

When we consider that there are different groups that have similar elements in each group, in a population we aim to reduce variance by drawing a separate sample from each of the groups and then combine the results to both reduce costs and also gather a representative sample. This sampling method is called stratified random sampling (Aczel, 2009).

An example for a stratification variable to reduce variance for in-depth accident data could be the time of the accident. The advantage of such stratification is that the costs are often lower and the error variance is lower while the principles are similar to random sampling. All of the groupings are required to be included in the sample, in this example the method has to cover all 24 hours in a day, and the population distribution would need to be known (Hagstroem et al., 2010).

2.9.3 Systematic sampling

In situations where a population is arranged in an orderly manner (e.g. goods in a warehouse) a random sample can be drawn in a systematic way compared to a simple random sample. To select a systematic sample of n elements from a population of N elements, we divide the N elements in the population into n groups of k elements (Aczel, 2009). Aczel (2009) states that "for this sampling method we randomly select the first element out of the first k elements in the population, and then we select every kth unit afterward until we have a sample of n elements" (p. 16-19).

2.9.4 Clustered sampling

When we do not know every element in a population but know that a cluster contains many of these elements we may choose to use the method of cluster sampling. This method can also be used if it is not feasible to sample the whole region but smaller sub regions are more easily sampled and a simple random sample or a stratified random sample may not be carried out as easily (Aczel. 2009).

If these clusters were considered as police zones for example a random selection of police zones could be carried out. It is important that all available zones are included not just those zones that are happy to participate, as this would compromise the findings in relation to the other districts. Attaining data using clustered samples can help reduce the distances that are travelled for data collection purposes, though in some cases this type of clustering increases rather than decreases error variance (Hagstroem et al., 2010).

2.9.5 Issues of sampling with real world data

Though theoretically these different types of sampling present solutions to many of the issues that arise due to sample representivity, it is necessary to take into consideration that these solutions will not be possible for all types of studies, particularly when considering time constraints and available funding.

The sampling considerations outlined above do not need to be considered with police reports, due to all accident data being present in the police data and the sample being representative of the population as a whole.

In-depth accident datasets have more difficulty in achieving this as the resources that they have are limited in nature. The collection of one case can lead to not collecting other available cases in the categorised area. A further difficulty is that sampling is not possible on a national scale with these types of studies, as an area that is deemed to be representative is selected and accident cases are collected from here, so the elements of randomness and generalizability are influenced.

In order to overcome these issues the FICA study concentrated on collecting a particular group of accidents to make it representative of all of these accidents in the system. While data collection in INRETS focuses on collecting as much information as possible on one case and then after collection of these cases are finished analysing the similarities and differences between these group of cases and police national accident reports.

2.10 Scenario development with real world data

Traffic accident scenario development concentrates on identifying causes and consequences that are common for a group of accidents. Traffic accident scenarios have been developed in one of two ways, the first way historically was for accident data (usually macroscopic data) to be interpreted by the accident investigator and expert in order to identify and relate similar accident types to develop scenarios.

Fell (1976) concentrated on a system that would explain human information processing failures or non-performance and tie it together with other factors. The accident causal system created causal chain links for human factors and chains for the other factors separately. Using this method as a base Malaterre (1990) analysed 72 in-depth accidents involving 115 road users collected as part of the INRETS study (stated in section 2.8) and identified 15 different scenario types in relation to this study using a method that focused on factors related to attention distribution, task competition and time pressure. This analysis was based on the causal charts that were filled in for this data and he tried to extend the causation charts to include information related to the emergency phase of a crash, but found that this was difficult to carry out and required further work. A further analysis of French data has been continuously carried out by the INRETS study using a similar methodology.

Historically in-depth accident data have not been used with statistical analysis that uses multivariate methods. The main reason for this is the small number of cases usually collected for these study types (often 50–70 cases) and their being a much larger number of descriptors for each specific case making it difficult to provide statistical weighting to carry out statistical procedures (Fleury & Brenac, 2001; Sandin & Ljung, 2007).

The second way to develop accident scenarios is to use macroscopic data to interpret accidents and create groupings using statistical analysis procedures. As the large number of crashes in national accident datasets make it suitable to use these methods, a number of studies have been carried out using this data to develop accident scenarios. With the introduction of computers that allow for the analysis of national data in an easier manner, researchers have used statistical programs to analyse national statistical accident data.

Skyving, Berg, and Laflamme (2009) and Berg, Gregersen, and Laflamme (2004) used two multivariate analysis techniques: the Factorial Analysis of Correspondence (FAC) and the Hierarchical Ascendant Classification (HAC) to identify scenarios for national accident statistic data with regards to fatal crashes to older drivers and road accidents occurring during driver training respectively, and each found 4 specific accident groups and developed countermeasure suggestions according to the specific factors targeted. They analysed data in terms of up to 15 subfields and 100 variables to understand how specific accidents occur.

Fault tree analysis methods have also been used to develop scenarios for accident types allowing for decisions based on different decision points to be made. This analysis has been used on both national statistical data (Chen et al., 2009) and in-depth accident data (de Oña, López, Mujalli, & Calvo, 2013a) to identify different types of accidents that occur with regards to injury outcomes and provide a discussion of countermeasures for these outcomes.

Latent class clustering methods have been used with national statistics data to demonstrate how different types of accidents occur (de Oña, López, Mujalli, & Calvo, 2013b; Depaire, Wets, & Vanhoof, 2008). These studies used multinomial logit estimation as a predictor variable, to quantify whether the accidents have a specific relation to the statistical analysis in real world data analysis, and Bayesian networks, to identify if the results that were found were new and if possible, insights could be gained. A discussion with regards to quantify countermeasure were also made in these studies.

2.11 Countermeasure analysis with real world data

Countermeasures have been developed historically to prevent accidents and accident injury. The aim of a countermeasure is to counteract risk. In road traffic, risk is a function of four elements (Porter, 2011):

- The exposure – the amount of movement or travel within the system by different users or a given population density

- The underlying probability of a crash, given a particular exposure
- The probability of injury
- The outcome of injury

The nature of the countermeasure changes according to which of these factors it is aiming to counteract. In trying to counteract the probability of a crash, active safety measures and interactive safety measures are used to prevent the road user from participating in an accident. To lessen the probability of injury passive and interactive safety measures are incorporated to lessen the severity of injury and also the consequences of the injury. To lessen the outcome of injury, rescue safety measures are incorporated to lessen the severity of injury and also the consequences of the injury. These preventions also aim to reduce the monetary costs related to an accident.

We can further divide countermeasures using the Haddon matrix to identify relevant groups. In terms of the human road user countermeasures can then be further developed based on different road user groups needs and structured accordingly for these needs. In terms of the vehicle and environment countermeasures can also be structured, depending on the different elements that are aimed to be altered.

The main factors that countermeasures aim to control are all the possible factors that cause accident risk to increase. For future improvement of road safety countermeasure approaches need to be altered. Countermeasures related to the human road user and aimed at the prevention of certain human behaviours can be universal in some senses, as human beings go through similar processes when in the traffic environment, though culture also need to be taken into consideration when applying certain countermeasures. When identifying countermeasures for the road infrastructure and environment the specific needs of the country and region need to be taken into consideration.

2.12 Effectiveness of safety measures

When analysing the role of the human user in the roadway, safety technologies aim at either providing support to the road user or undertaking a behaviour in place of the road user. Intelligent transport safety functions must not only be adapted to drivers needs but also be restricted in order not to overload or disturb drivers' capacity.

Intelligent Transport Systems (ITS) are a term for any electronic, information processing, communication and control technologies that may be used in the transport domain (Bayly, Fildes, Regan, & Young, 2007). Bayly et al. (2007) identified ITS systems based on the most common classification of the system as either;

- In vehicle based
- Infrastructure based
- Cooperative

In vehicle based systems either provide information to the vehicle user, automate some form of the driving behaviour or intervene in a vehicle user's behaviour and adapt it or stop it from occurring. Infrastructure based systems provide roadside messages or user information gathered from the road users to control traffic flow, and cooperative measures involve communication between the different systems within the traffic system either between vehicles or with the infrastructure. This information can either be one way or two way (Bayly et al., 2007). When a road user is faced with a potential conflict situation the amount of time available to make a critical decision is relatively short and for these systems to be used effectively fast interpretation of the drivers needs are necessary.

Atalar et al. (2012) identified two steps of analysis necessary to assess the potential effectiveness of a safety system:

1. First the capacity of the system to correctly address drivers' needs has to be estimated by comparing the functionalities of the system with the difficulties met by the driver in the accident situation. This asks for a clear and precise description of the way the system is acting.

2. Then it must be taken into account the physical and operational constraints found in accident situations that the system shall be able to compensate for, in order to be fully efficient. This necessitates a thorough understanding of the specifications of the system functionality.

In terms of how ITS effects safety Draskóczy, Carsten, and Kulmala (1998) identified 10 items that are necessary to be used as a barometer in terms of safety;

1. Direct effects of an in-car system on the user (modification of the driving task)
2. Direct effects of a road-side system on the user
3. Indirect, behaviour modifying effects of the system on the user
4. Indirect, behaviour modifying effects of the system on the non-user (imitating effect)
5. Modification of interaction between users and non-users (including vulnerable road users)
6. Modifying accident consequences (e.g. by improving rescue, etc.)
7. Modifying exposure (frequency and/or length of travel)
8. Modifying modal choice
9. Modifying route choice
10. Modifying speed choice

2.13 Road user accident statistics

The nature of an accident determines both the injury outcome and the possible countermeasure that can be used to prevent this type of accident. In order to understand the types of accidents that occur, it is necessary to identify when and where specific types of traffic accidents occur. An analysis of the accident configuration would in turn allow for more detailed and focused countermeasures to be developed. According to an overview of STATS19 national data from the Great Britain in 2007 (DfT, 2013) the most common crash types that occur are:

- Rear-end crashes
- Overtaking crashes
- Crashes at intersections/Turning crashes
- Single vehicle crashes

A rear-end crash refers to a crash in which the front of one vehicle collides with the rear of another vehicle (Singh, 2003). Rear-end crashes are the most common type of crashes that occur. Data compiled in the United States by the National Highway Traffic Safety Administration (NHTSA), found that approximately 29.7% of all crashes in the year 2000 were rear-end crashes. These crashes were responsible for 30% of all injuries and 29.7% of the property damage accounted for by all accidents in that year (Singh, 2003). A way a driver can limit the possibility of a rear-end crash is by maintaining a distance from other road users that is appropriate for the driving conditions. A proper space cushion can be defined as that which provides a driver adequate time to recognise a potential hazard and make a decision to avoid this hazard by potentially bringing the vehicle to a stop (Abdel-Aty & Abdelwahab, 2004).

Kuge, Ueno, Ichikawa, and Ochiai (1995) identified behavioural issues with regards to rear-end accident situations, by observing road users driving in the road environment, as:

- Highly dangerous situations were more frequently observed in the vehicle speed range over 100 km/h.
- In approaching a preceding vehicle in motion, the driver of the following vehicle does not expect the preceding vehicle's emergency braking when judging the brake timing.
- In approaching a stopped vehicle, the higher the approaching speed is, the less the time allowance becomes, resulting in higher deceleration than normal.

Davis and Swenson (2006) reviewed rear-end accident video data and using simulations identified three possible causes of rear-end collisions as (1) too short following distances were probable causal factors for the collisions, (2) one or more of the vehicles ahead had a longer reaction time than the preceding vehicle, and (3) had the reaction times been equal the crash probably would not have occurred.

According to Clarke, Forsyth, and Wright (1998) overtaking accidents accounted for 7.9% of fatal road accidents in the county of Nottinghamshire, England, between the years 1989–1992, and the percentage of cases with serious injuries was over 20%. Clarke et al. (1998) analysed 100 overtaking accidents from national police data compiled in the Nottinghamshire region within the UK and found 10 specific scenarios for these accident types. The three scenarios with the largest number of accidents were (1) accident collisions with a right-turning vehicle either due to a young driver makes a faulty overt, or an older driver making a faulty right turn, (2) a head-on collision, and (3) a loss of control accident which is particularly significant for younger drivers.

Clarke et al. (2005) also carried out an analysis of police report files to identify what type of accidents occurred most frequently at junctions. The drivers that were over-represented were the youngest and oldest groups of drivers, and they were the least likely to stop before turning. The young drivers particularly had problems when turning onto major roads. With regards to gender differences women were more likely than men to stop before turning, tended to have their collisions with other women and were under-represented as drivers of the non-turning vehicle (Clarke et al., 2005).

Sandin (2009) analysed causation charts for 52 drivers involved in 26 in-depth investigated urban intersection crashes in Sweden using the DREAM method. The aggregated charts identified six risk situations, four for drivers without the right of way and two for those with the right of way. In two risk situations, one for drivers with and one for the drivers without the right of way, common patterns showed that the drivers had not seen the other vehicle due to distractions and/or sight obstructions. For drivers with right of way a common pattern was that they did not expect another vehicle to cross their path and so did not take this into consideration. Though the roadway in Sweden is not the same as the UK, some of the basic features of intersection crashes may be similar due to the psychological processes that are being undertaken in these situations being similar for road users.

A prominent example of a traffic manoeuvre that leads to an accident in intersections is the ‘looked but did not see’ accidents. ‘Looked but did not see’ accidents can be defined as accidents where the road user looks in the direction of the other vehicle but does not see or perceive the presence of the other road user (Herslund & Jørgensen, 2003). A main reason of these accidents is thought to be experience, as less experienced drivers are expected to make these types of functional failures more than drivers that have more experience as they are considered to be better at identifying `danger` (Koustanai et al., 2008). The nature of the driving situations determines the cognitive load that is necessary, less cars and less cluttered environmental scenery will allow for the road user to identify situations more easily and give road users more time to react to the potential functional failure situation and in turn cause a near miss rather than an accident (Underwood et al., 2003).

For example, to overtake a vehicle, drivers have to estimate both the time interval in relation to that vehicle and the risk of collision with vehicles coming in front of them, it is very difficult for them to alter this manoeuvre once they have started it, while other manoeuvres have lower cognitive loads for the drivers (Clarke et al., 1998; Koustanaï et al., 2008; Summala, 1996). Thus road users develop cognitive schemas and behavioural steps for different manoeuvres. Depending on how demanding the workload is the road user can adapt to the situation or their late response, or no response, can be the cause of a functional failure occurring. For situations that occur as expected it is not difficult for the driver to react appropriately, but when a driver is faced with a situation that is novel and not expected the cognitive workload becomes a constraint that may limit the road user correctly interpreting what the other road users are going to do as well as how the environment is going to change.

Koustanai et al. (2008) compared different participants using simulations in terms of hazardous situations while overtaking a vehicle or while turning left when the road user could either predict danger or they cannot predict it. They found that in situations where drivers cannot predict change in the behaviour of a vehicle that is overtaking, there is a much higher level of accidents occurring compared to situations where they cannot predict danger with left hand turns.

Brorsson, Rydgren, and Ifver (1993) analysed a sample of questionnaires from 467 (62% response rate) injured single vehicle occupants in crashes in Sweden, calculating 95% confidence intervals that take into consideration the random variability and comparing them against the collected data. They identified that the risk level of men between the ages of 18-19 years of age were eight times greater than the risk level of men 25-54 years of age. Within this sample one third of all drivers were suspected to be drunk, and this expectation was more common in middle aged drivers compared to younger drivers.

Clarke et al. (2007) analysing fatal accidents in the UK using a qualitative assessment method of police reported data found that nearly 40% of the sample as a whole were single vehicle accidents. Of the accidents that resulted in fatalities thirty six per cent were single vehicle accidents. The number of fatal injury cases where drivers were considered primarily at fault also had a higher number of single vehicle accidents (47% of all fatal injury cases) and disproportionally skewed towards involving drivers between the ages of 17-20.

Analysing single vehicle accidents from insurance company data gathered in Finland between the years 1978-1991, Laapotti & Keskinen (1998) found that male and female drivers involved in loss of control accidents were similar in terms of proportion. The configuration of these accidents were different as male drivers loss of control accidents were usually single vehicle accidents, while female drivers loss of control accidents resulted in multiple vehicle accidents. Male driver’s loss of control accident contributing factors were speed and alcohol, and typical occurred during the evening or night and female drivers loss of control accidents contributory factor were typically slippery roads.

Sandin and Ljung (2007) analysed 38 cases of in-depth single vehicle accidents in Sweden and demonstrated four scenarios:

- Vehicles drifting out of their lane due to going off in a certain direction and lacking in recovery attempts. When fatigue was a contributory factor these accidents occurred mostly on high speed roads and where distraction was a factor these accidents occurred on lower speed roads.
- Loss of control in curves with reduced road friction due to the roads slipperiness from roadway factors.
- Excessive speed in curves where the drivers realised too late that they were approaching the curve at too high a speed.
- Alarmed drivers reacting with excessive manoeuvres due to the other vehicle drifting toward there lane.

2.14 Vulnerable road user accidents

Vulnerable road users are identified by the WHO as making up 46% of all global fatalities in road safety throughout the world (WHO, 2009). Vulnerable road users consist of pedestrians, powered two wheeler riders and young/elderly persons. Due to their high attribution of traffic accidents young drivers are also considered as vulnerable road users in some literature.

2.14.1 Powered two wheeler (PTW)

Powered two wheeler (PTW) riders are one of the most at risk user groups within the traffic environment. Statistical data show that each year they represent 15% of people killed on European roads, and according to the World Health Organization nearly 200,000 deaths in the world (WHO, 2006).

The nature of PTW accidents is different from other vehicles in the traffic environment due to both the physical dimensions as well as the more limited safety measures that can be implemented, as PTWs have a higher degree of manoeuvrability compared to other vehicle types.

Previous research has demonstrated that PTW riders are the most at risk group of vehicle riders/drivers within the road traffic environment. PTW rider death rates as a function of distance travelled are generally found to be about 30 times greater than for car occupants (Johnston, Brooks, & Savage, 2008). In the United Kingdom in 2010 there were over 403 motorcycle riders (including moped riders) killed in road crashes, 4780 killed or seriously injured (KSI) and over 18,686 involved in recorded injury crashes (all severities) (DfT, 2010). In 2009 there were 140 deaths and 1,709 people killed and seriously injured (KSI) per billion vehicle miles for motorcycle riders. The corresponding figures for car drivers were 3 killed and 30 KSI per billion vehicle miles (DfT, 2010). Huang, Siddiqui & Abdel-Aty (2011) using US police accident data gathered in the state of Florida between the years 2000-2007 reported that the odds of PTW riders being injured are 2.63 times higher than for drivers of light vehicles. The PTW rider was reported as not at fault in 43% of these accidents (Haque, Chin, & Huang, 2009).

A number of studies on powered two wheelers have identified that males are part of the accident population on 85% or more of these accidents (Bjørnskau, Nævestad, & Akhtar, 2012; MAIDS, 2009). Yannis, Golias, and Papadimitriou (2005) identified that rider age was a significant factor in the causation of a motorcycle accident as crash involvement decreased with increasing driver age. Engine size was also identified in increasing the severity of the accident but not affecting the possibility of causation. A number of studies have identified that rider age, alcohol impairment, speed, rider attention, road surface and road class all influenced accident severity (Chorlton, Conner, & Jamson, 2012; Preusser, Williams, & Ulmer, 1995; Shankar & Mannering, 1996).

With regards to motorcyclists being at fault Haque et al. (2009) found that a number of factors were more likely in the motorcyclist being considered at fault, these factors are (1) cases on motorways or high speed limit roads, (2) when the motorcycle engine size was larger, (3) slippery roads, (4) intersection conflicts, and (5) when the motorcycle rider was either an older or younger rider. Motorcyclists were less likely to be at fault during crashes that occurred at night time or at locations where surveillance cameras were present (Haque et al., 2009). Seiniger, Schröter, and Gail (2012) found that younger motorcyclists are more likely to be at-fault in the event of a collision, as are riders who are under the influence of alcohol. Similarly, motorcyclists were less likely to be at-fault when the other driver was of younger age or was driving under the influence of alcohol. Clarke, Ward, Bartle, and Truman (2004) when asking motorcycle riders to fill in self-report forms and state who were at fault from motorcycle accidents had a result of car user 78% of the time.

The literature dealing with motorcycle safety has highlighted one of the main type of accident as situations where a motorcycle rider having priority on a straight road is put in conflict with another road user when this road user moves in front of the rider despite not having priority (Clarke, Ward, Bartle, & Truman, 2007; Hurt, Ouellet, & Thom, 1981; MAIDS, 2009; Peek-Asa & Kraus, 1996; Williams & Hoffmann, 1979; Wulf, Hancock, & Rahimi, 1989). Furthermore, these accidents appear to be characterised by an often high level of injury severity (Pai & Saleh, 2008; Pai, 2009; Peek-Asa & Kraus, 1996). Williams & Hoffman (1979) identified motorcycle visibility as the prime cause in 64.5% of motorcycle to car accidents, with particular importance being placed on the front of the motorcycle.

Clarke et al. (2007) identified accidents occurring on bends as some of the most dangerous accidents with double the risk of rider or passenger fatality. These types of accidents were found to be mainly caused by the motorcyclists and the rider was found to have a very high likelihood of being inexperienced (Clarke et al., 2007).

In an in-depth study of motorcycle accidents Clarke et al. (2004) identified the three most common types of motorcycle accident scenarios. The most prevalent scenario was a right of way violation where a vehicle pulled out from a side road onto a main carriageway into the path of an approaching motorcycle. In these accidents the driver in conflict with the motorcycle rider typically reports detection issues despite feeling that they have satisfactorily scanned the roadway before turning. This has been termed a ‘look but failed to see’ error (Brown, 2002). Clabaux et al. (2012) analysed a small sample of motorcycle crashes in an in-depth manner in France and found that in urban areas “looked but did not see” accidents involving motorcyclists are related to initial speeds (for the motorcyclists) that are significantly higher compared with those of other types of accidents in intersections, but there was no difference in rural accident speeds. The high frequency of junction accidents involving motorcycles, and specifically right of way violation accidents, has also been reported.

One of the main difficulties for a driver in identifying an approaching motorcycle is that the driver either cannot correctly identify the speed that the motorcycle is travelling at, if at all. Crundall, Crundall, Clarke, and Shahar (2012) performed two experiments using video clips to understand the situations in which car drivers were less likely to see motorcycle riders than other vehicles. These situations were similar to ‘look but failed to see’ accident types. Drivers who were also motorcycle users were more cautious compared to both experienced and novice drivers, with novice drivers performing the worst.

Preusser et al. (1995) uncovered five specific crash types for PTW riders in the US that were, (1) ran off-road (41%), (2) ran traffic control (18%), (3) oncoming or head-on (11%), (4) left-turn oncoming (8%), and (5) motorcyclist down (7%). Left turns and failure to yield were common factors associated with the involvement of other motorists in these accidents.

2.14.2 Pedestrian accidents

Pedestrian accidents are a particular group of accidents in which the road user is not protected during a traffic accident, by either a structure or protective clothing. Rather the immediate collision is with the other object in the road, and thus the injury outcomes of these accidents are usually greater than other accident types. Also, due to the injury outcomes of pedestrian accidents being more severe than other types of vehicle accidents, pedestrians are more dependent on other road users’ behaviours and adherence to traffic laws.

Road traffic crashes involving pedestrians often occur and include a large proportion of all fatal and serious injury accidents (WHO, 2009). With regards to fatal injuries resulting from traffic accidents close to 50% involve vulnerable road users (WHO, 2009). Approximately 21% of road traffic deaths involve children, yielding an average of 720 child deaths related to road traffic accidents per day.

In the Netherlands almost 50% of pedestrians involved in fatal accidents are over 65 years old, and the group that has the next highest proportion are children under the age of 14 (Hummel, 1998). In the U.S., fatal injuries are more likely for males of all ages and they account for 70% of pedestrian deaths. The fatality rate per 100,000 population was 2.19 for males compared to 0.91 for females (Clifton & Livi, 2005).

In the United States, in 2007 approximately 73% of pedestrian fatalities occurred in urban areas, this is thought to be due to the larger amount of pedestrian activity in urban areas (NHTSA., 2008). Research in Europe also has a higher rate of pedestrian fatalities occurring in urban areas (SafetyNet, 2009). Despite a larger number of pedestrian accidents occurring in urban areas the probability of a fatal injury occurring in a pedestrian accident is 2.3 times more likely in rural areas compared to urban areas (Mueller, Rivara, & Bergman, 1998).

Hunter, Stutts, Pein, and Cox (1995) analysing approximately 5,000 pedestrian crashes (and 3000 bicycle crashes) from six US states found that the most common pedestrian accident types were dart-out in first half of the street (24%), intersection dash (13%), dart-out in second half of the street (10%), midblock dart (8%), walking along roadway (7.4%), and turning-vehicle (5%) accidents.

Analysing naturalistic driving data Habibovic and Davidsson (2012) identified two main causation patterns in car user to pedestrian accidents. The first pattern occurred when drivers did not identify that a pedestrian was present in an intersection, due to either visual obstructions and/or because they were concentrating on another aspect of the road environment. In incidents away from intersections, the above defined situation also occurred as well as situations where pedestrians unexpected behaviours led to conflicts.

Also, although it is not a specific crash type, approximately two thirds of pedestrian fatalities throughout the world occur at night or under low-light conditions. The other road users are not able to see during the night and this is a contributory factor to the accident types provided above (Zegeer & Bushell, 2012).

A driver speeding in the roadway can increase the risk to pedestrians in several ways. First, vehicle stopping distance increases substantially as vehicle speed increases. Second, the risk of a pedestrian death from a collision with a motor vehicle is much greater for higher vehicle speeds (Zegeer & Bushell, 2012). According to a study by the UK Department for Transport (2010), the probability of pedestrian death is 85% when the striking vehicle is traveling at 40 mph. This probability drops to about 45% for a 30 mph impact and drops further to 5% if the vehicle is traveling at 20 mph at impact.

In a study by Tefft (2013) analysing US NASS pedestrian crash study data between the years 1994 and 1998, results show that the average risk of a struck pedestrian sustaining an injury of Abbreviated Injury Scale 4 or greater severity reaches 10% at an impact speed of 17.1 miles per hour (mph), 25% at 24.9 mph, 50% at 33.0 mph, 75% at 40.8 mph, and 90% at 48.1 mph. The average risk of death reaches 10% at an impact speed of 24.1 mph, 25% at 32.5 mph, 50% at 40.6 mph, 75% at 48.0 mph, and 90% at 54.6 mph. The difference between risks according to age should also be considered as the risk of death for a pedestrian that is 70 years of age was similar to the risk for a 30-year-old pedestrian struck at a speed 11.8 mph faster.

2.15 The scope and aim of this thesis

In the first section of this chapter an explanation of how drivers acquire driving skills was outlined (Fitts & Posner, 1967). From the literature it was understood that different models (descriptive or functional models) provide different interpretation possibilities when analysing data and understanding how the road user understands the traffic environment is paramount when considering human error. Human error models (Norman, 1981; Rasmussen, 1982; Reason, 1990) were also considered. In order for individual cognitive issues to be understood in the realm of traffic safety the necessity for a model based approach was underlined in the literature.

Previous studies (Sabey & Staughton, 1975; Treat et al., 1979) provided a detailed explanation with regards to the causes of traffic accidents for previous years, new technological developments within vehicles requires new solution measures to be provided.

A review of studies that have identified accident scenarios identified that recent in-depth accident studies (INRETS; GIDAS; FICA) have concentrated on obtaining detailed data on accident causation and have used causation charts as a way of understanding how accidents occur. As there is only a small number of accident cases for each chart the implications are more limited than compared to national data statistical studies that provide scenarios.

The limitations of the national accident data scenario studies (Bédard et al., 2002; Skyving et al., 2009) are that the level of detail that these studies obtain is not sufficient for thorough accident countermeasure development purposes. Only a basic understanding of the differences between accidents is available from these analysis types.

Aggregating accident data has been identified as a tool to understand how necessary countermeasures can be highlighted though causation data has not been previously used with statistical methods for this process as the small number of cases collected did not allow for this type of analysis.

Though a wealth of detail is provided for different studies the necessity for in-depth data that provides volatile and stable data as well as interviews and analysis of the driver's behaviour was underlined. Though national data and other data sources provide insight with regards to how certain behaviours occur, a more detailed systems approach is necessary for a clearer understanding of driver errors and related factors.

A systems theory approach needs to be applied to analyse a combination of factors that cause risk in the traffic environment. This is especially more important with the implementation of ADAS and infrastructure related changes.

This PhD project aims at creating a methodology to analyse accident datasets by statistically analysing different accident scenarios based on specific factors and manoeuvre types. The later research focuses on risk factors, causes and scenarios. This will be done by identifying the main functional failure (perceptive failure) that a road user makes and identifying the other factors that contributed to the accident occurring. A holistic methodology to analyse these accidents is used. This analysis allowed an identification of specific factors that occur in different accident scenarios, their frequency and stage at which these factors occur.

2.15.1 Research questions

The research questions that this PhD aims to answer are:

- **What are the most frequently occurring traffic accident scenarios in relation to driver error (for each accident)?**

  An identification of similar accident scenarios will allow for factors that are closely related to be examined. This question will be answered by analysing accident cases collected during the OTS study using a multivariate statistical methodology to identify multiple factors and their relationships within the accidents.

- **What interactions occur for two vulnerable road user groups (pedestrians and powered two wheeler riders) in relation to driver error on the part of both road users?**

  This question will be answered by analysing pedestrian and PTW crashes collected from the OTS dataset.

- **Are there any differences in establishing causation factors between microscopic and macroscopic data?**

- An analysis of in-depth (on the spot data) and national data (Stats19 data) will be carried out and a comparison of the level of detail that they provide will be made in chapter 7
- **How does the interaction between human failure and contributory factors cause traffic accidents?**

  Will be analysed by understanding how the statistical methodology allowed for these interactions to be grouped as different accident scenarios. The interactions within different scenarios will be explained and discussed with regards to other research.

2.15.2 Research objectives

The objectives of this PhD project are,

1. To develop an analysis method that will allow for statistical analysis to be carried out on causation sequence chains in large traffic accident datasets.
2. To analyse all relevant accident scenarios in the OTS dataset.
3. To analyse the causal chains to understand how functional failure sequences occur within particular accident groups to develop accident scenarios.
4. To understand the links between interacting factors and individuals to further understand how these interactions cause accidents to occur.
5. To identify countermeasures implications for the scenarios that are highlighted in the research with regards to different stakeholders in the road traffic environment.

2.16 Summary

Chapter 2 presented prominent theories and models in relation to the acquisition of driving skills, analysis of human, vehicular and infrastructural factors to traffic accidents, accident causation studies and road user behaviour studies. An explanation of the cognitive processes that road users carry out in learning how to drive and possible models to help understand these processes with regards to accident analysis was made.

The different causes of traffic accidents were then identified with a description of how drivers acquired driving skills. A discussion of human error in terms of a number of relevant models was carried out. Historical approaches to accident investigation and different countermeasures used during these periods were discussed.

Different real world studies that aimed to collect data in terms of road use behaviours were identified and described. An explanation of sampling necessities was carried out as well as a discussion of some of the limitations of sampling when conducting an in-depth study. Scenario development and countermeasure development with regards to real world data were also discussed. A brief literature review of common road user accident types and configurations as well as a separate consideration or powered two wheeler riders and pedestrians was further carried out in this chapter. Finally a description of the limitations of the past research and the questions and aims of this thesis were presented.

3 Classification of road user error

3.1 Selection of accident causation method

The reviewed literature identified the importance of analysing traffic accidents using a multifactorial based model and a data collection process that would gather all relevant data. The study carried out in this chapter included a review of three current accident causation analysis methods used throughout Europe that identify and classify human error/driver behaviour as a basis of accident causation interpretation and adhere to a time based structure when coding for each road user in a specific accident. The aim of this review was to select a method that would enable collected traffic accident cases to be coded with regards to all applicable contributory factors and human failure in a clear manner.

These three methods aim to identify the nature of a crash by identifying all relevant information from the crash on site within a timeframe that allows for the perishable data from the crash to be obtained.

This chapter will present a comparative case study with regards to these three causative analysis methods;

1. Driver Reliability and Error Analysis Method (DREAM)
2. Accident Causation Analysis with Seven Steps (ACASS)
3. Human Functional Failure (HFF)

This comparison will consist of three separate sections. The first section compares the main failures and contributory factors identified by each classification system. The differences between the coding systems are explained as well as the possible different interpretations that would arise as a result of the accident cases being coded with the different methods.

The second section consists of a questionnaire study comparing the three methodologies with regards to questions on the methods applicability and usability. This questionnaire was filled in by ten participants from different accident research centres spread throughout Europe. An inter-rater reliability comparison between the three methods was carried out with six participants, using five cases from three different countries.

In the third section a comprehensive comparison of the human failures coded for each accident as well as all human, vehicular, environmental and infrastructural factors was carried out. This study also analysed whether different conclusions and countermeasures would be identified as a result of using the different methods by comparing the results identified by each coding method. The questions that this study answered are as below;

- Form a better understanding of how human failure is coded by each method (section 3.3)
- Identify how coding is similar/differs from one method to another (section 3.5)
- Compare the usability of the methods (section 3.6)
- Compare the inter-rater reliability of the methods (section 3.7)
- Analyse whether different interpretations for integrated safety measures would result from different analysis for each case (section 3.8)
- Identify the most suitable method in regards to identifying specific human failure (section 3.9)
- Identify the method most appropriate for handling OTS data (section 3.9.6)

3.2 Review of recent accident causation models

Understanding how and why traffic accidents occur is necessary for the correct implementation of road safety measures. Past research in this field has concentrated on developing passive safety measures and improving vehicle structures so that the consequences of crashes would lessen. These measures concentrate on the issues that a road user faces after a crash occurs, namely injuries and other consequences of the crash. Present research is also concentrated on active and integrated safety measures that alleviate issues that would/may lead to a traffic accident.

The methodologies analysed here work under the assumption that if human cognition and different types of human error could be clearly understood, an identification of the necessary countermeasures would be possible. The aim of accident causation research is to identify relevant factors and combine them with the main failure that the road user made to better understand how the accident occurred.

The three accident causation analysis methods that have been reviewed use causal accident theory as a basis, and integrate systems theory and behavioural theory methods to analyse both different risk factors and how the traffic environment interact with road users in specific situations.

3.3 Recent accident causation models

In this section a brief description of three of the more prevalent accident causation models that have been developed, and are currently in use by in-depth accident investigation teams throughout Europe, will be made.

3.3.1 Driving Reliability and Error Analysis Method

The Driving Reliability and Error Analysis Method is based on the Cognitive Reliability and Error Analysis Method (CREAM) developed by Erik Hollnagel (Hollnagel, 1998). CREAM was originally developed to analyse safety-critical incidents in nuclear power plants (Sagberg, 2007). CREAM is based on the “man-technology-organisation” (MTO) perspective and categorizes causation to cover these three perspectives, as well as a method to identify the relationship between the categories, causal factors and error modes (Sagberg, 2007).

The aim of DREAM is to systematically explain the causes of road accidents using in-depth investigations as the basis of classifying these accidents. The goal of DREAM is to identify traffic situations where the development and introduction of technical solutions would decrease future accidents (Wallén

Warner, Björklund, Johansson, Ljung, & Sandin, 2008). In order to do this DREAM concentrates on identifying interactive systems for risk avoidance, to allow the driver to work within the system to limit the number of dangerous situations that they are confronted with (Wallén Warner et al., 2008).

DREAM is made up of three main elements: (1) an accident model, (2) classification scheme, and (3) a method to acquire all the necessary data to be acquired. DREAM uses common performance conditions (CPC) parameters forms which cover two main areas: (1) the general driving conditions at the scene, and (2) the driver's general condition which allows the researcher to frame the accident situation (Wallén Warner et al., 2008). The CPC form is used to identify and collect all relevant data in the accident site in order to be able to codify the accident into DREAM.

The next step of DREAM is to organise the contributing factors and to connect them. In order to do this the analyst uses empirical data and their pre-understanding of accidents to determine a causal relationship that brought the accident to fruition. Empirical data are classified as phenotypes (or critical events), they are the observable consequences of the accident and are situations that hold true for nearly every accident that occurs, so in order to be able to classify accidents it is necessary to use them to form groups. DREAM defines the phenotypes as limited to the time, space and energy continuum. Phenotypes are objective and use general and specific phenotypes as the groups, the investigator using general phenotypes if there is limited information and specific phenotypes if there is all necessary information.

Genotypes classify the causal factors preceding the critical event. They usually have to be deduced by the accident investigator from the scene, information from drivers and witnesses, so they are wholly subjective. Genotypes are the MTO perspective re-adapted for use within an accident model. Man is the road user, technology is the vehicle and organisation is broken into two groups, infrastructure and organisation corresponding to the road environment and traffic management (Ljung, 2007). The subgroups below these groups are the basis for the DREAM analysis and allow the investigator to interpret the necessary links to develop an understanding of how the accident occurred. An extended example of the coding of an example case for genotypes and phenotypes can be seen in figure 6.

Example explanation (adapted from Ljung, 2007): This example is of an accident from Sweden and as such concerns a vehicle travelling on the right hand side using a left hand drive steering system. In this example driver A approaches a T junction that the driver is familiar with intending to go straight ahead. As driver A’s vision was blocked by a hedge, driver A could not see driver B. When driver B pulls out driver A does not have any time to make an avoidance manoeuvre and so drives into vehicle B’s left side. Thus for this reason the phenotype selected is timing: no action. The general genotype selected combined to the phenotype is misjudgement of the situation as driver A thought the intersection was free to enter when it was not. For the lower level genotypes missed observation (the hedge obstructed driver A’s view), inattention (as a result of being familiar with the junction) and expectance of certain behaviours (driver A expected the other driver to stop on the T junction in accordance with traffic laws) are chosen. At this level the analysis is stopped as no other genotypes can be selected. An illustration of the accident and the DREAM causation chart that were attributed can be seen in figure 5 and figure 6.

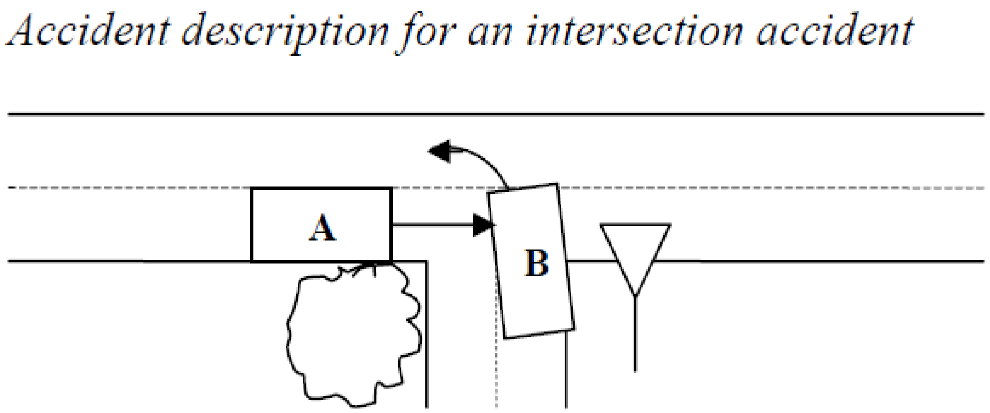


Figure 5: Illustration of the example accident (Adapted from Ljung, 2007)

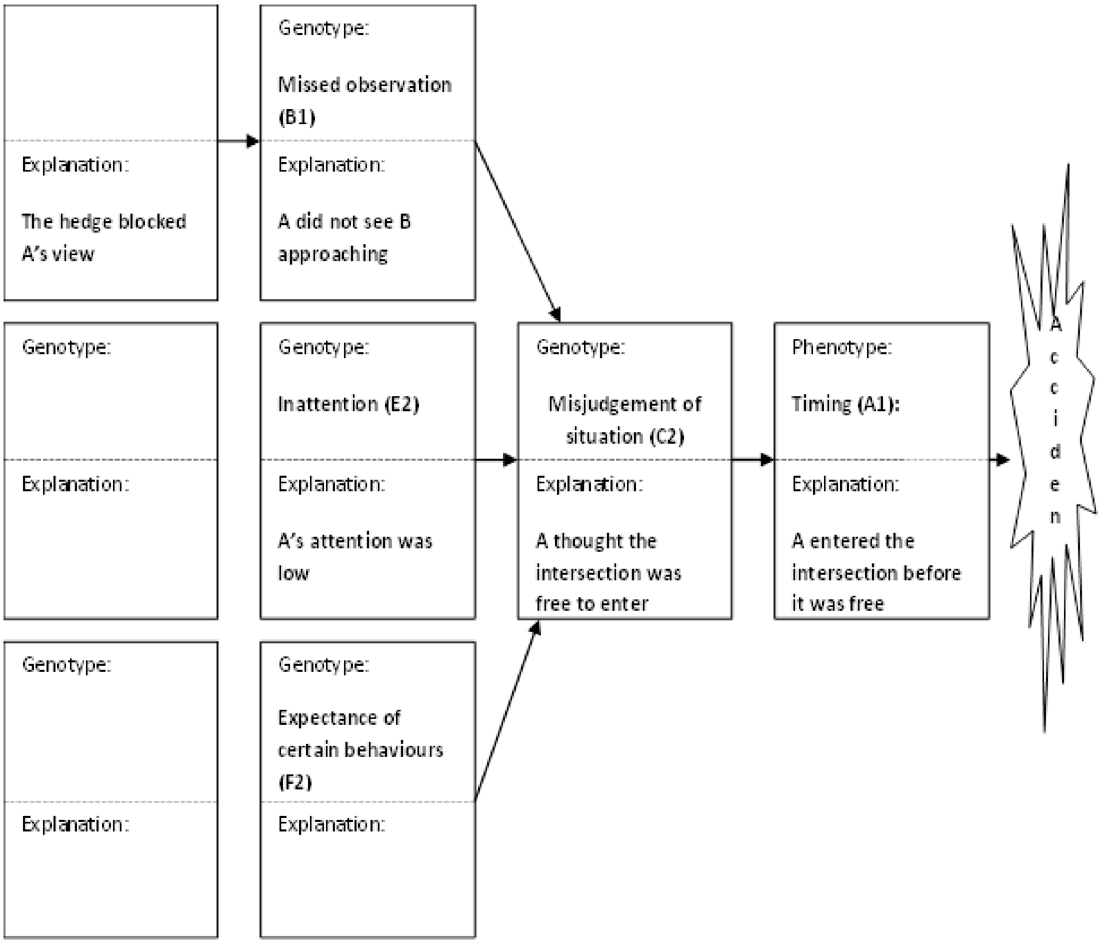


Figure 6: DREAM causation chart (Adapted from Ljung, 2007)

3.3.2 Accident Causation Analysis with Seven Steps method

The Accident Causation Analysis with Seven Steps (ACASS) method places particular emphasis based on the psychological and perceptual aspects when interpreting accidents. This is an accident causation system that aims to describe relevant human causes of traffic accidents. The analysis aims to start at the level of human functions and processes before the traffic accident happens, identifying initial conditions, perception, judgement and acting (action) leading up to a crash. The levels of perception are presented as a causal chain link, taking steps to reach the next level and all in all coming to seven steps.

The first level is the driver's perception of the situation which is broken down into three sublevels: visibility, observation and recognition (of the situation).

The second level is the assessment of the accident which has two sublevels, evaluation and planning, the driver's evaluation if there is a hazardous situation or if no action needs to be taken. The next level is the action level which is broken down into two sublevels, the selection and execution stages, which relates to which action was selected from the plan, if it was correctly selected and the execution of the action which leads to the accident or near miss situation (Pund, Otte, & Jaensch, 2006).

When the information is presented to the driver it is first perceived (step 1), then observed (step 2), recognised (step 3) and evaluated (step 4). Then plans are put forward about how to deal with this information (step 5), an appropriate plan is selected (step 6) and put into operation (step 7) (Otte, Jaensch, & Pund, 2007). In the context of implementing ACASS into non-psychologist based on scene data collection it appeared to be sensible to simplify the seven categories of human causation factors, to improve the practicability of this system during on scene investigations for team members. Thus two changes were performed:

1. The categories (2) observation and (3) recognition were merged to one category information access.
2. The category (6) selection was merged into the category (7) operation.

The remaining five categories (figure 7) are the main categories of human causation factors and may easily be converted back into a seven step system with the knowledge of the specific influence criteria of the categories. Figure 7 demonstrates this conversion (Otte, Jaensch & Pund, 2007).

As an addition to this there are also human factors (symptoms of a disease, age risks, intoxication with substances and individual risk factors), technical factors from the vehicle (technical defects, maintenance failures/ condition of the vehicle, human/machine interface and vehicle design) and factors from the environment and infrastructure (conditions, road design, factors from nature, other external influences and roadside objects) which can all be added as causative factors of road accidents. These factors help the investigator understand the other factors that contributed to the human causes of the accident.

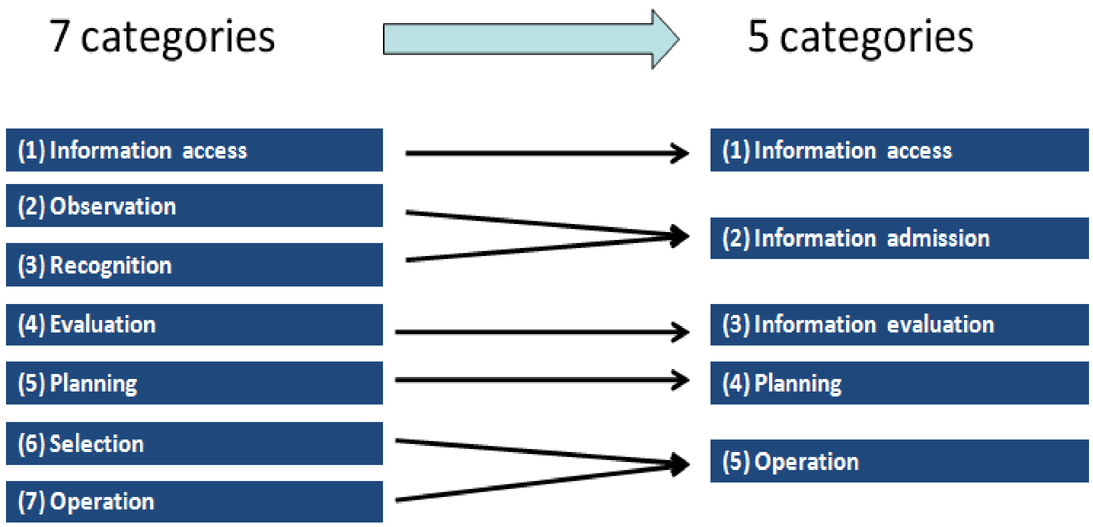


Figure 7: Seven steps categories converted to five categories (Adapted from Otte, Jaensch & Pund, 2007).

3.3.3 Human Functional Failure method

The HFF method uses a human factors approach to categorize the factors and situations within an individual accident in such a way that a non-human factors expert will find useful. The HFF method uses a holistic approach with the fundamental ergonomic model as a basis. The user (road user), task (driving, walking, and running) and tool (the vehicle) are related to each other and the environment (road user's surroundings) surrounds the task. Using this foundation pre-accident driving situations were determined that relate to the task (driving) and factors which lead to the contributing factors and main failure that is related to the driver, vehicle and the environment (Naing, Bayer, Van Elslande, & Fouquet, 2007).The pre-accident driving situations are divided into two levels, the task and location. The task is further divided into three: primary level driving tasks (essential to the journey), secondary level driving tasks (important to the journey but not essential) and tertiary level driving tasks (not directly related to the journey). The location is related to where the vehicle is leading up to or during the accident as different situations necessitates different reactions, for example if the vehicle was at an intersection this is different compared to a single lane road. HFF divides the sequence of events in an accident into four phases, connected to one another (Molinero et al., 2008):

1. The driving phase: the driving situation can be described as the one in which the user is before a problem arises. It is the 'normal' situation, which is characterised for the driver by the performance of a specific task in a given context, with certain objectives, certain expectations, and so on. It is 'normal' because no unexpected demands are made upon him.
2. The rupture phase: the 'rupture' is an unexpected event that interrupts the driving situation by upsetting its balance and thus endangering the system.
3. The emergency phase: it is the period during which the driver tries to return to the normal situation by carrying out an emergency manoeuvre.
4. The crash phase: the crash phase comprises the crash and its consequences.

Factors are determined for the user, vehicle and environment. The user’s state, experience and behaviour are described in terms of the accident, and the environment, road condition, geometry, traffic condition, visibility, traffic guidance and other factors are considered as possible risk factors. For the vehicle mechanical function, maintenance, design and load are classified as separate factors which are possible causal factors of traffic accidents (Naing et al., 2007). The interactions within this system are illustrated in figure 8. Using these levels of analysis a broad understanding of the accident is accounted for and thirty commonly occurring scenarios were also identified using French data in order to be able to specify and group accidents and determine the main faults of the accidents and developing countermeasures for these scenarios.

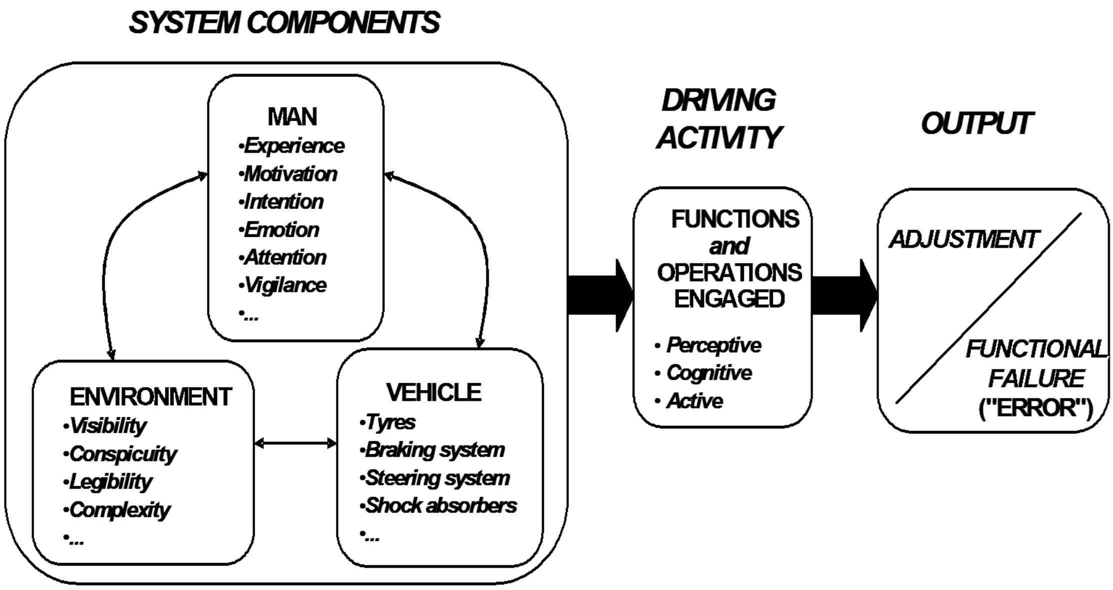


Figure 8: Interactions within the elementary human-vehicle-environment system (Adapted from Van Elslande & Fouquet, 2007)

3.4 Accident causation European studies

The three methods stated above have been used in a variety of projects conducted throughout Europe as part of the 6th and 7th EU Framework Programme for Research and Technological Development (FP6 & FP7). An explanation of three projects that were carried as part of the above stated framework programmes are carried out below.

SafetyNet

The SafetyNet study was a sixth framework European Union funded project aimed at the development of a new European Road Safety Observatory (ERSO) to gather data and knowledge to inform future safety policies. The SafetyNet Accident Causation Database was developed between 2005 and 2008. It contains in-depth data on 997 accidents covering all injury severities, collected from accidents that occurred in Germany, Italy, The Netherlands, Finland, Sweden, and the UK. The data was collected 'at scene' or 'nearly at-scene' and complemented by follow up interviews, using a common methodology across all countries. Causation data was recorded according to the DREAM methodology that was renamed as the SafetyNet Accident

Causation System (SNACS) methodology. A case study using both DREAM and ACASS coding methods was also carried out during this project.

TRACE

The TRACE project (Traffic Accident Causation in Europe) was a sixth framework European Union funded project aimed at developing a scientific accident analysis at identifying, characterising and quantifying the nature of risk factors, groups at risk, specific conflict driving situations, and accident situations. Estimations of the safety benefits of a selection of technology-based safety functions were also provided by the TRACE project. This project used HFF as its methodology with regards to accident causation.

DaCoTA

The DaCoTA project was a seventh framework European Union funded project aimed to further extend and develop the ERSO, by developing and implementing new approaches to gather, structure and apply policy-related safety data. Work package 2 was interested in developing a Pan-European In-depth Accident Investigation Network and reviewed the three different accident causation methods explained in this chapter with regards to usability throughout all of Europe.

3.5 Comparison of methods

The three methods described above are all based on an underlying cognitive model that allows for a time-based analysis of traffic accidents. DREAM uses a model based on latent failure conditions and makes a distinction between sharp end and blunt end failures (Ljung, 2007). A traffic accident or incident is caused by the failure of a joint driver-vehicle-system at a certain point in time and space (a failure at the sharp end), and the analysis of the event needs to find out which factors generated this sharp end failure. It must also determine if any blunt end failures contributed to the development of the event. A blunt end failure is a failure, which can be remote in space and time, but the consequences of which still can be an important contributor to the course of the event.

ACASS identifies the basic disturbance in the respective step of the hierarchically structured flow chart, describing the human basic function in detail. This function is perceived as an error during the process of the information processing and action conversion. The failure of a basic human function is explained due to effective physiological or psychological factors, e.g. perception errors due to distraction, decision errors due to unsolvable conflicting objectives or action errors due to coordination errors. The role of the motivation of the drivers concerns (above all the risk evaluation of a situation and the driver's behaviour) concerning the motivational conditions, particularly in the steps "estimate" (interpretation of the recognized characteristics) and "planning" (action draft due to intention formation) are asked.

The HFF method uses a model based on research carried out by Michon (1985), Rasmussen (1982) and Reason (1990) for the analysis of traffic accidents. Errors, slips and violations are defined in accordance with different accident causation factors at the pre-crash stage.

All three models require the failure to be put in the context of contributing variables (risk factors) that trigger the potential of a functional failure occurring, combining these factors identifying a causal sequence of events. In DREAM the critical (physical) events that occur to cause the accident are coded as part of the causation chain link, though the cognitive stage that the road user was at is not coded. Despite this difference in coding, the underlying accident process is coded in a manner that is similar to the other two analysis methods, with regards to coding contributory factors.

Contributory factors are divided into three specific groups:

1. Human factors
2. Vehicular factors
3. Environmental/Infrastructure factors

The DREAM method only allows certain links to be made between the contributory factors and the critical event. A study was carried out to identify links between contributory factors and critical events based on previous accident research studies. Any previous chain links that were possible with previous versions of DREAM were replaced if not present in the literature (Warner & Sandin, 2010). Nevertheless a large number of connections can still be made using this model, though some of the factors are needed to be used as a linkage factor to go in certain routes.

The ACASS method requires each accident chain sequence to be coded using a predetermined structure. Only particular paths can be chosen for each specific failure and factor outcome. Though more than one chain can be coded, the specific structure of the chain is determined by the factors that are included. Vehicle, environment and infrastructure factors can be selected as necessary.

The HFF method allows the coder to select all applicable contributory factors separately to the functional failure. For the functional failures, certain scenarios have been developed from work that has been carried out in France by IFSTTAR, though these are not necessary to be coded unless applicable, to allow the researcher to code specific failures with more detail.

3.5.1 Example case for the three methods

A case example from the UK OTS database was coded using the three methods in order to demonstrate the different coding styles and possibilities.

Case example: In this example the driver is driving along a B class road with a 60 Mph speed limit with a slight bend to the left hand side. The driver is confronted with an animal in the middle of the driving lane and has swerved to avoid hitting it. The driver has swerved to the offside and lost control. The vehicle has left the road to the offside and collided with the ditch to the offside. The vehicle has then rolled before coming to a rest.

DREAM coding: The phenotype selected for this case is too short distance, as there was not enough distance for the driver to make a reactionary behaviour. The genotypes selected were misjudgement of time gaps and priority error.

ACASS coding: The main failure selected was planning and the type of human factor was decision error. The environmental factor run off the bend was included in the chain.

HFF coding: The main failure selected was a prognosis 3 failure, which is the driver driving in such a way as to not be able to react to any obstacles the driver is confronted with. The driver was identified as driving at an inappropriate legal speed and having newly obtained a driving license, and the car as a newly purchased car. The bend was also coded as an environmental factor.

3.5.2 Comparison rationale

The comparison conducted aimed to identify the method that was most suitable for use in this thesis. An understanding of the cognitive stages that a road user goes through is necessary in order to understand the type of active safety measures that are needed to combat the safety issues that road users are faced with. This study aims to take cases that have been coded within an in-depth accident database and use different models of accident causation to analyse these cases. For this purpose the analysis performed had four main purposes. These purposes were to compare;

1. The ease of use of each of the methods as an analysis tool of traffic accidents.
2. The ease of comparison of the methods and use as a wide spread analysis tool by more than one accident research centre.
3. The differences between the methods with regards to a preliminary understanding of any differences between the human failures that were identified and also other factors relevant to the accidents occurrence.
4. The ease of identifying all relevant failures and factors in a traffic accident, and using this data for analysis purposes.

This comparison was carried out as part of the DaCoTA project that ran between the years 2010 to 2012, and aimed to increase awareness with regards to important issues in traffic safety using a scientific basis to further knowledge on road safety issues. Naturalistic driving data and in-depth accident data were also collected, and policy issues were analysed within this project.

The data used in this study was collected by DaCoTA work package 2, which was interested in developing a Pan-European In-depth Accident Investigation Network. Studies 1 and 2 were carried out by the work package 2 partners as part of the selection procedure for an accident causation method for the DaCoTA database and the data was collected by the author. The analysis carried out in this chapter was conducted separately from the DaCoTA analysis.

3.6 Comparison of usability

3.6.1 Objective

The objective of this comparison was to identify the usability of the three accident causation methods described above according to experienced accident researchers throughout Europe. A questionnaire was designed to allow the participants to rate each method with regards to four separate areas:

1. The ease of use
2. Inter-rater reliability
3. Description of the accident
4. The accident outputs usability

3.6.2 Participants

This study sent separate questionnaires for each of the three methods and had a total of nine responses for the DREAM and ACASS questionnaire and eight responses for the HFF questionnaire. Each questionnaire had the same questions and format.

The participants were from nine different centres that were experienced in in-depth accident research. An experienced researcher was identified as any researcher that had carried out at least thirty or more in-depth accident investigation on scene or retrospectively.

Three of the participants were from the centres that developed the accident causation models Chalmers, Sweden (DREAM), Medical University of Hannover, Germany (ACASS) and IFSTTAR, France (HFF). The remaining participants were from the Transport Safety Research Centre (United Kingdom), IDIADA (Spain), Hellenic Institute of Transport (Greece), SWOV Institute for Road Safety Research (the Netherlands), General Directorate of Traffic (SPAIN) and Cidaut (Spain). A distribution of the different investigators that participated in the questionnaire is identified in table 6. All of the accident investigators had investigated at least 30 cases, either on scene or retrospectively. Prior experience with the different methods were evenly balanced out with one investigator having prior experience with all of the methods, three investigators having experience with two of the methods, two investigators having experience with one of the methods and three investigators having no prior experience with any of the methods.

Table 6: Accident investigator experience and prior knowledge

Accident investigator number	Gender	Age group	In-depth investigation experience	Prior experience with the methods		
				DREAM	HFF	ACASS
1	Male	26-35	Over 300 cases	Yes	Yes	No
2	Male	36-45	Over 500 cases	Yes	No	Yes
3	Male	36-45	Over 200 cases	Yes	No	Yes
4	Female	26-35	Over 100 cases	No	Yes	No
5	Male	26-35	Over 30 cases	Yes	Yes	Yes
6	Female	36-45	Over 100 cases	No	Yes	No
7	Male	46-55	Over 50 cases	No	No	No
8	Male	36-45	Over 200 cases	No	No	No
9	Male	36-45	Over 100 cases	No	No	No

3.6.3 Procedure

The questionnaires were compared with regards to 4 different closed (yes/no) questions with multiple sections and 1 open ended question. The questions were:

Closed questions

- How easy is the coding system to learn?
- Would you expect/can you demonstrate good inter-rater reliability?
- Does this coding system allow you to fully describe all aspects of any accident?
- Does the output fully explain the cases?

Open ended question

- On average how long did it take you to code using each system?

In total the questionnaire included 25 questions in these groups with “yes” and “no” alternatives. The questionnaire also included a question asking the amount of time taken to code the cases with each separate method. Nine accident researchers in total filled in the questionnaire. Participants did not answer some of the questions and this resulted in a number of questions having less than nine responses.

3.6.4 Results

Table 7 illustrates the results that respondents gave with regards to the question “How easy is the coding system to learn?”. Of the coders 50% had previous experience with the DREAM method, 33% with ACASS and 44% with HFF. The user manual was referred to the most for DREAM (62.5%), then HFF (37.5%) and then ACASS (11.1%). ACASS (88.8%) was rated most often as a fairly intuitive method, compared to DREAM (62.5%), HFF (11.1%) score was particularly low for this item. DREAM was rated as requiring specialist knowledge by 63% of respondents, compared to HFF (37.5%) and ACASS (11%), which was only rated with a positive response for this item by one of the respondents.

Table 7: How easy is the coding system to learn?

Question	DREAM		ACASS		HFF	
	Yes	No	Yes	No	Yes	No
Prior experience with method	4	4	3	6	4	5
Manual is not needed to refer to	3	5	8	1	1	8
System fairly intuitive	5	3	8	1	4	5
Specialist knowledge is not required	3	5	8	1	5	3
Sufficient coding possibilities	5	3	9	0	6	2
Clear start and end	5	2	8	1	9	0
Benefit from further training	4	4	3	6	8	1
Total	29	26	47	16	37	22
Percentage (%)	52.7	47.3	74.6	25.4	62.7	37.3

ACASS was rated as having sufficient coding possibilities by all of the respondents, while DREAM (37.5%) and HFF (25%) were rated by some of the coders as having too many possibilities. The start and end of all of the methods was identifiable by most of the respondent's, DREAM (71.4%), HFF (100%) and ACASS (88.8%). Most respondents rated that they would benefit from further training in HFF (88.8%), this items rating was 50% for DREAM and 33% for ACASS.

Table 8 illustrates the results that respondents gave with regards to the question "Would you expect/can you demonstrate good inter-rater reliability?". For DREAM (12.5%) and ACASS (11.1%) only 1 of the respondents felt that the codes were not clear with regards to coding. While for HFF, only 3 of the 7 of the respondents felt that the coding choice was clear. For ACASS all respondents felt that the coding choices were not difficult to make while for DREAM (37.5%) and for HFF (62.5%) respondents felt that that the coding choices were more difficult to interpret and make. Most respondents felt that there were more than one interpretation for the codes with regards to DREAM (62.5%) and HFF (72.7%), while this was less than half of the respondents with regards to ACASS (42.7%).

Table 8: Would you expect/can you demonstrate good inter-rater reliability?

Question	DREAM		ACASS		HFF	
	Yes	No	Yes	No	Yes	No
Meaning of each code clear	7	1	8	1	4	3
Coding choices easy to make	5	3	9	0	3	5
One interpretation for coding	3	5	4	3	2	5
Sufficient factors to choose	8	0	8	0	8	0
Similar coding expected regardless of coders background	3	5	7	2	0	8
Total	26	14	36	6	17	21
Percentage (%)	65.0	35.0	85.7	14.3	44.7	55.3

None of the respondents considered that any of the methods had too many factors to choose from. Interpretation based difficulties were expected for HFF (100%), DREAM (62.5%) and ACASS (28.6%) depending on the coders theoretical background.

Table 9: Does this coding system allow you to fully describe all aspects of any accident?

Question	DREAM		ACASS		HFF	
	Yes	No	Yes	No	Yes	No
Contains enough relevant factors	5	2	2	7	8	0
Can code all factors needed	3	4	3	5	5	3
Involves a time sequence	4	4	1	7	9	0
Includes all involved users	7	0	4	4	9	0
Suitable for simple/complex cases	7	1	3	6	9	0
Total	26	11	13	29	40	3
Percentage (%)	70.3	29.7	31.0	69.0	93.0	7.0

Table 9 illustrates the results that respondents gave with regards to the question “Does this coding system allow you to fully describe all aspects of any accident?". Of the three methods HFF was the only method deemed to contain all relevant factors, 71% of respondents felt that DREAM contained enough relevant factors and 28% felt that ACASS contained enough relevant factors. Respondents felt that the systems did not allow all necessary factors to be coded, for DREAM (57.1%), ACASS (62.5%) and HFF (37.5%). All respondents felt that HFF contained a time sequence for events, 50% felt that for DREAM and 11% for ACASS. They felt unanimously that all involved users could be coded for HFF and DREAM, while this was 50% for ACASS. For the question "Analysis was suitable for complex cases" for HFF the response was 100%, 85% for DREAM and 33% for ACASS.

Table 10: Does the output fully explain the cases?

Question	DREAM		ACASS		HFF	
	Yes	No	Yes	No	Yes	No
Output provides clear contributory factors/causes	8	0	7	2	9	0
There are variables that are not reflected in output	4	3	4	5	8	0
Systems output is manageable	7	1	9	0	8	1
Suitable for single and aggregate analysis	4	1	6	0	7	0
Can answer key research questions	4	1	1	4	6	0
Helps develop and identify countermeasures	7	0	6	3	8	0
Total	26	6	26	12	37	1
Percentage (%)	81.3	18.8	68.4	31.6	97.4	2.6

Table 10 illustrates the results that respondents gave with regards to the question "Does the output fully explain the cases?". All respondents identified that HFF identified all contributory factors and reflected all variables in the output. All felt that DREAM and ACASS identified a clear cause but did not allow for all factors to be coded. Respondents felt that the output was manageable for ACASS (100%), HFF (88.8%) and DREAM (86.5%). All respondents felt that HFF was suitable for single case and aggregate analysis, to answer key research questions, and to help identify and develop countermeasures. Most respondents felt that DREAM (80%) could answer key research questions, though this was not the case for ACASS (20%).

Table 11: On average how long did it take you to code using each system?

Coding system	Prior experience	N	No prior experience	N	All respondents	N
ACASS	7 minutes	2	14 minutes	6	12 Minutes	8
HFF	33 minutes	4	30 minutes	4	32 Minutes	8
DREAM	18 minutes	3	38 minutes	4	25 minutes	7

Table 11 illustrates the average time that respondents took to code each accident case with the different accident causation methods. The respondents were required to answer this question by only taking into account the average time that it took for a case to be coded using the accident causation methods once the case had been fully reviewed. ACASS was coded the fastest, by both respondents with no previous experience and prior experience on average taking 12 minutes to code a case. HFF took an average of 32 minutes to code and there was no difference between coders with prior and no prior experience. DREAM took an average of 25 minutes to code and there was an average difference of 20 minutes between coders with no prior experience and those with prior experience.

Figure 9 demonstrates the total accumulative percentages that the different methods received as a composite criterion from all of the closed items in the questionnaire study. These percentages demonstrate that the HFF method had a higher percentage of suitability (73.6%) according to the questionnaire study that was carried out compared with DREAM (65.2%) and ACASS (66.7%).

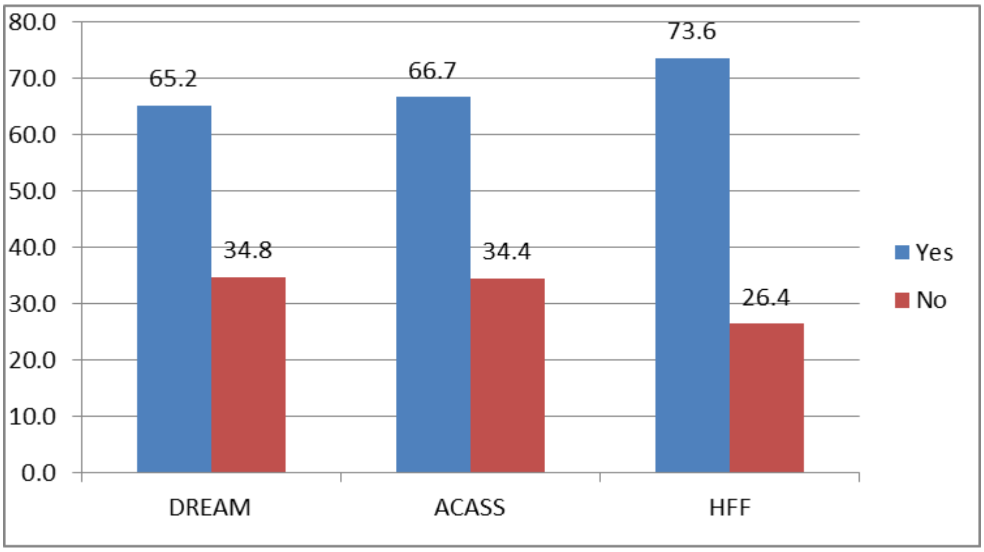


Figure 9: Total scores (Accumulated from tables 3, 4, 5 and 6) by percentage

3.7 Inter-coder coding comparison

This study analysed and compared the inter-rater reliability of the three accident causation methods. This study was collected as part of the DaCoTA work package 2 accident causation comparison study.

3.7.1 Participants

Six researchers from six different traffic research institutes throughout Europe were asked to code five cases provided by three different accident research centres. The cases were then sent for comparison in terms of inter-rater reliability. All participants had experience using at least two of the methods and were experienced accident analysts.

The participants were based in the TSRC in Loughborough University, United Kingdom, SAFER in Chalmers University, Sweden, MUH in the University of Hannover, Germany, CIDUAT in Spain and SWOV in the Netherlands. IFSTTAR based in France only contributed codes for the HFF methodology.

3.7.2 Procedure

Two accident cases were selected from the Intact database (Sweden), two accident cases were selected from the IFSTTAR database (France) and one accident case was selected from the GIDAS database (Germany). Each accident was presented with a detailed description of the accident, multiple photos of the accident scene and road users involved, and a diagram of how the accident occurred. These cases were provided by different centres so differences in the detail level of the accidents were present, though each case provided a detailed and adequate amount of information.

Five separate analysts analysed each of the cases using DREAM and ACASS and six analysts analysed the cases using HFF. Each of the analysts were advised to select the appropriate codes for each accident, as the codes were not already made. The case coding was carried out by the accident investigators using the data provided.

As the HFF method requires a full coding of the accident only the accident causation section was taken for this comparison. The coding comparison was divided into three sections; (1) the main failure, (2) human factors, and (3) vehicular factors and environmental/infrastructural factors. As the DREAM method codes the type of phenotype (the critical event) rather than the failure, the phenotype was used in place of the main failure for comparison purposes. Also, due to the HFF method having a more thorough definition of a failure the failure type was taken into consideration rather than the specific failure groupings.

For the ACASS method it is possible to select more than one failure. For this study the first failure that was coded was selected, and for the second or third failure only the contributory factors were included. The ACASS and HFF method allows road users that were non-active in the accident to be coded in the analysis. For the ACASS method the road user was coded as being non-active and no coding required eight times, and for HFF this situation was coded four times. For this study only factors that were determined as definitely contributing to the accident were compared, so some factors that were coded as possibly contributing were not included in the analysis.

The analysis was carried out using Krippendorf's Alpha. This method is suitable for inter-rater reliability studies where there are more than one individual rating the cases factors that are present, and can also compare comparisons when data is missing. This analysis was run using an SPSS macro that was developed by Hayes (2005). For this study as the amount of coded variables differed, only lines that had at least 2 or more of the analysts' codes were used. So for example if three participants coded 1 failure, 2 human factors and 1 other factor, 1 participant coded 1 failure, 3 human factors and 1 other factor and 1 participant coded 1 failure, 4 human factors and 1 other factor, the comparison would be conducted on 1 failure, 3 human factors and 1 other factor.

3.7.3 Results

The total number of road users that were coded in this study can be seen in Table 12. One of the coders did not code for 2 of the HFF cases thus the final number of coding for the HFF method was 56, 4 of these road users were identified as being passive in the accident, and as the method allows this, were not coded and thus were excluded from this analysis. For the DREAM and ACASS methods five coders coded 10 road users, though for the ACASS method 8 of the cases were described as passive cases and no functional failure was coded.

Table 12: Total number of users coded

Method	Total road users coded	Total road users not coded	Total number
DREAM	50	0	50
ACASS	42	8	50
HFF	52	4	56

The total number of contributory factors coded can be seen in figure 10 for each of the three methodologies. DREAM had the most codes with a total of 133 of which 115 concern human factors, HFF had 100 contributory factor codes with 77 of them being human factor codes and ACASS had 97 codes with 85 of them being human factor codes.

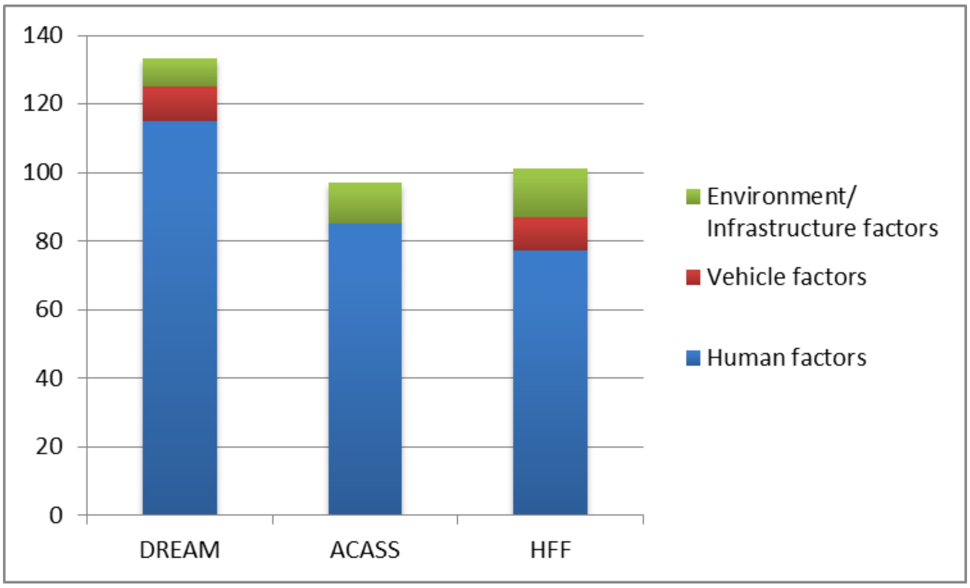


Figure 10: Total number of contributory factors

Table 13 illustrates the results from the Krippendorff's Alpha comparison done for each of the three accident causation methods with regards to inter-rater reliability. The results identified that out of the three methods the DREAM method had the highest rating of inter-rater reliability in both the main failure (.655) and the human factors coding (.514). HFF had the second highest inter-rater reliability ratings for the main failure (.471) and the lowest for human factors (.270). ACASS had the lowest inter-rater reliability rating for the main failure (.318) and second highest rating for human factors (.336). The other factor codes were highest for HFF (.322) compared to DREAM (.193) and ACASS (.133).

Table 13: Inter-rater reliability ratings with Krippendorff's Alpha

Method	Main failure match	Human factors match	Other factors match
DREAM	.655	.514	.193
ACASS	.318	.336	.133
HFF	.471	.270	.322

3.8 Case coding comparison

This study measured similarities in the final causation factors attributed to a set of accidents when coded to each of the three methods.

3.8.1 Participants

For this study three analysts that had experience in using each of the three methods were selected to analyse 23 accident cases. The participants were researchers in the Transport Safety Research Centre, Loughborough. The analyses of the cases with DREAM were taken from the SafetyNet study. The analyses of the cases with HFF were taken from the TRACE study. The analysis of the cases with ACASS was carried out separately.

3.8.2 Procedure

This study used a total of 23 cases, 18 cases were taken from the On the Spot study database (United Kingdom), 2 cases from the Intact database (Sweden), 2 cases from the IFSTTAR database (France) and 1 case from the GIDAS database (Germany). These cases were in-depth accident cases which included relevant information with regards to the human, vehicular and environmental/infrastructure factors that contributed to the accident occurring. Relevant scene measurements, photos, videos and interview data was obtained for each of the accidents and analysed. For each of the cases one analyst that had experience in using the different accident causation methods analysed the data using one of the three different methodologies. Each of the analysts were advised to select the appropriate codes for each accident, as the codes were not already made. Each coding was made from the available data provided to the accident coders.

3.8.3 Comparison of methodologies

This study used a three level comparison with regards to the crashes to examine human failure pre and post-crash factors. A differentiation was made between the human, vehicular and environmental/infrastructure factors that contributed to the occurrence of the crash. A separate comparison was conducted for the main failure that was outlined by the method in order to compare the perceptual stage that related to crash behaviour.

Due to the fact that the HFF method requires coding for the whole accident, it was determined that in the name of comparison only two of the coding steps the “pre-accident situation” and “initiating factors” would be used for this test of coding comparison. Due to the difference in the structure of the methods, the failure that was coded and the number of factors coded were compared. This comparison will be used to determine whether all of the coding methodologies would result in a similar suggestion of a solution or whether they would be different.

At the first level the number of pre-crash contributory factors that were coded for each coding method was compared. These were divided into three categories: human, vehicular and environmental/infrastructural factors. The second level used the same comparison structure but analysed factors that occurred during the accident. The main analysis was a comparison of the stage of failure that was coded. As DREAM uses 4, ACASS uses 5 and HFF uses 6 levels of cognition perception during the understanding of traffic behaviour the comparison was conducted as below.

Due to these differences an identification approach was applied to the cases. For each accident case, it was determined whether the codes ‘exactly matched’, were ‘similar’ or were ‘not a match’. Table 14 makes a brief comparison of the different methods with regards to how they code human failure, their objectives and the cognitive model number of subgroups present in each.

Table 14: Comparison of Accident Causation methods

Method	Main failure	Failure types	Method objective	Cognitive model
DREAM	8 Subgroups	18	Identifying intelligent systems for risk avoidance	4 sub-groups
ACASS	5 sub-groups	17	To compile an evaluation - neutral coding system of accident causes	5 sub-groups
HFF	6 sub-groups	30	To propose countermeasures well fitted to the real needs of road users	6 sub-groups

3.8.4 Results

Table 15 demonstrates the matches between the DREAM coding and the other two methods. DREAM was an exact or similar match with ACASS on 54% of the cases, when taking the cases that have not been coded into account, and with HFF on 67% of the cases.

Table 15: Matching between DREAM and other methods

Method		ACASS	HFF	Total
DREAM	Exact Match	21	22	43
	Similar	6	8	14
	Not Matching	23	15	38
	Not Coded	2	7	9
	Total cases	52	52	104

Table 16 demonstrates the matches between the ACASS coding and the other two methods. ACASS was an exact or similar match with DREAM on 54% of the cases, when taking the cases that have not been coded into account, and with HFF on 58% of the cases.

Table 16: Matching between ACASS and other methods

Method		DREAM	HFF	Total
	Exact Match	21	13	34
	Similar	6	13	22
ACASS	Not Matching	23	19	42
	Not Coded	2	7	9
	Total cases	52	52	104

Table 17 demonstrates the matches between the HFF coding and the other two methods. HFF was an exact or similar match with ACASS on 58% of the cases, when taking the cases that have not been coded into account, and with DREAM on 67% of the cases.

Table 17: Matching between HFF and other methods

Method		ACASS	DREAM	Total
	Exact Match	13	22	35
	Similar	13	8	21
HFF	Not Matching	19	15	34
	Not Coded	7	7	14
	Total cases	52	52	104

3.9 Discussion

3.9.1 Study results

In the questionnaire results ACASS was identified as the easiest method to learn, and code in a timely manner. HFF was identified as being the most difficult to understand, and most respondents identified that they would prefer more training with this method if possible. Out of the three coding methods DREAM was described as being the least intuitive to code, HFF and DREAM were both identified as having more relevant factors and being better suited for developing countermeasures by the participants, with the HFF method having a slight advantage over the DREAM method.

When analysing the composite scores from the comparison the HFF method gathered more positive scores compared to the other methods. HFF was identified as coding the most satisfactory number of factors, allowing the most thorough analysis of the cases, and having the highest possibility of countermeasure identification. The HFF coding procedure provided the most contributory factors and failure types for the accident investigator to utilise, and so allowed for a larger differentiation between the accidents.

The results of the inter-rater reliability analysis highlighted that the DREAM method had the most matches with regards to both the main functional failure and human contributory factors. The DREAM method has a more rigid structure with regards to coding and only allows for certain chains to be developed. The method provides the most constraints in terms of accident analysis, though if selecting a method that aims at providing similar codes was particularly important the DREAM method would be more suitable than the other two methods.

A similar study using the DREAM method, that analysed inter-rater reliability of nine participants on 4 specific case types, also identified that when users who have been trained in DREAM have coded less than 5 cases the results differ for the majority of the coding, compared to similar coding by users that have coded more than 5 cases (Warner & Sandin, 2010). In this study genotypes, which can be called specific causation or contributory factors, average agreement of 83% and phenotypes, which can be called single (main causation factors), average agreement of 78% was observed (Warner & Sandin, 2010). This study prepared 3 cases as training cases for the participants, and once the participants had coded and sent their codes the solutions were sent to the participants, after which the study was carried out. Participants that had coded less than 5 cases previous to the study were left out of the study.

Both DREAM and ACASS require a higher level of abstraction to make sure that the groupings of the coded cases are similar to each other and provide similar results. The links that can be made between factors and failures are constrained for both of these methods. Though this allows for a clearer link to be made between the cases some of the level of detail within the cases may be lost as a result.

In terms of the similarities between coded cases HFF and DREAM had a 67% match which was significantly higher than ACASS with DREAM (54%) or ACASS with HFF (58%). Exact replication of coding for the different methods is difficult for a number of reasons. As seen in the inter-rater reliability comparison even when the same cases and methods are given to individuals unless they have extensive training in making their codes uniform in nature, there will be a large number of differences.

It should be also taken into account that only one coder for each of these three methods was used in this analysis and a further analysis with multiple coders that have been trained and have used each of these particular methods may have yielded different results.

3.9.2 How applicable are these methods to use with UK data?

When using a method that has been developed in another country for another culture the question of suitability is bound to be raised, more so with regards to traffic accident data as the environmental and infrastructure aspect of the data is certain to be different. One of the issues is the difference in terms of sampling, as these methods have been developed in other countries sampling crashes that have been used in other countries rather than in the UK. The INRETS studies where the HFF method was used, had a non-random sampling procedure selecting French accidents based in the Salon de Provence region (Morris, Smith, Chambers, & Thomas, 2005).The FICA study in which the DREAM method was used concentrated on single vehicle and intersection crashes, and the sampling plan was non-representative (Ljung Aust, 2010). The GIDAS study where ACASS was used had a random sampling plan based on representivity of national data is used (Otte, Jaensch, & Pund, 2007). The methodology that each centre used for data collection has similarities, but there were also differences due to the sampling criteria differences outlined above.

The TRACE and SafetyNet studies reviewed data from different European countries and developed all three coding systems to be applicable in all of these countries. The DaCoTA study aimed to overcome these differences by proposing a uniform manner of data collection and analysis procedures as well as sampling technique uniformality.

Furthermore, an initiative to provide in-depth accident data throughout Europe has been carried out by DaCoTA and both a case study and test trial of a database was carried out within that research project. The work package was tasked with selecting a method that was suitable to analyse crash causation data for a European consortium of road users between DREAM, ACASS and HFF.

The aim of the work package with regards to selecting the accident causation scheme was to select a method that provided (Hill et al., 2012);

- Good inter-coder reliability
- Possibility to make single case analyses and automated aggregated analyses
- Have a theoretically established background
- Sufficient number of relevant causation factors
- Clearly described contribution factors/causes
- A manual including examples and recommended applications
- Clear start and end points in the crash sequence

- Identification of the users of the data
- Results to suggest countermeasures
- Database implementation of all involved road users
- Some kind of time sequence

The selection procedure was made by the partners using the five example cases coded in this chapter and filling out the questionnaire that was coded in this thesis. After this process a voting procedure for each member to select their preferred methods in rank order were carried out. The team ultimately selected the DREAM method to be used in future research as it was determined that this method both allowed for relatively high inter-rater reliability and an analysis of the sequences of crash causation accidents to analyse this data with regards to both countermeasures and active safety components. According to Hill et al. (2012) all three steps in the process carried out in the work package showed a small advantage for the DREAM method.

The DREAM method developed in SafetyNet and during this project obtained support as the European method by the Commission. Furthermore DREAM was built into the database that DaCoTA was using and the use of another method would have been difficult. The considerable time necessary to transfer the other methods onto the database and the benefits of DREAM in terms of inter-rater reliability ratings meant DREAM was chosen for DaCoTA.

3.9.3 Model comparison in relation to a Safe System approach

When comparing the three different accident causation models it is important to underline that system management should minimise and be resilient to human error. A Safe System approach to road safety requires that system design, operation and management be taken into consideration when using accident coding and analysis methodologies. The aim ultimately is to understand how the system can better cope with the requirements that are placed on the functions.

This in turn leads to a number of latent factors being needed to be taken into consideration when looking at road safety requirements. The models that have been considered aim to find the cause of accidents based on understanding the factors that can be measured on the scene or through an interview process/use of questionnaire data. The purpose of understanding human error is to relate it to the accident situation, rather than consider the human to be at fault. The different accident causation methods aim to relate the conditions of the crash and the human error made to allow for the relevant accident sequence to be understood.

If a Safe System approach were to be used a better understanding of higher level factors related to how the road rules and regulations have been developed, and the effect that these factors have on individual's performance is necessary. This is taken into account by the DREAM method as the contributing level of different road users is not differentiated when coding an accident, but a further development of the methods by taking into account all relevant higher level system based factors would be beneficial.

All of the accident causation coding methods compared, concentrated on gathering information that was observable within the accident site and the subjective factors that could be obtained through interpretation of the incident and interviews. This allowed for a high level of information related to the incident to be collected, but did not provide an understanding of all relevant latent factors that are related to the incidents occurrence on a higher level.

The understanding of unobservable latent factors related to crash causation is an important issue, as the methods that are currently available require a level of subjective analysis to piece the factors of a collision together. All of the accident causation methods outlined aim to bridge the gap between a description of the internal conditions of the road user and the external conditions that the road user is faced with within the time span that leads to a crash (Van Elslande & Fouquet, 2007). The models selected focused on analysing the complex nature of interactions leading up to and during each individual reported incident.

The inclusion within the different methods of higher level factors related to design, operation and management issues would help clarify latent factors within the crash sequence that are currently unobservable. These issues are particularly important when finding solutions to issues with regards to what road users saw and why they carried out the behaviour. The understanding of these interactions and latent factors would provide a more detailed understanding of human error.

Despite providing relevant information with regards to each individual accident, the consideration of the other levels of traffic safety (e.g. government, local authority, management, front line operation) as outlined by the Safe System approach needs to be expanded upon (Salmon, Lenné, Stanton, Jenkins, & Walker, 2010). This would require the implementation of country specific model development, understanding the exact latent factors that are present within the traffic environment based on country specific policy and procedure. This could be possible by merging the above methods with system based approaches (Rasmussen & Svedung. I., 2000; Reason, 1990; Reason, 2000) to further identify the latent factors that are present, combined with the factors already coded for.

3.9.4 Discussion outcome

Due to the nature of the research that is being carried out in this thesis it was necessary that the accident causation analysis method chosen provide the most amount of flexibility in order for the statistical methods used to be able to demonstrate specific scenario selections. As one of the aims of the thesis was also to statistically identify significant accident clusters as well as develop countermeasures the last two questions in the questionnaire study were of a particular importance ("Does this coding system allow you to fully describe all aspects of any accident?" and "Does the output fully explain the cases?"). In the answer to both of these questions the HFF method was the most popular according to this group of researchers.

In cases where a large number of users code cases then DREAM would be a more suitable method as the inter-rater reliability would be higher compared to the other two methods, but since only one coder would be coding the cases for this analysis the inter-rater reliability was less relevant. HFF was deemed to provide better flexibility in relation to coding compared to both DREAM and ACASS and was selected to be used in the data analysis.

When interpreting accident data, it is important to be able to immediately understand what is happening, but also necessary to understand the latent factors that are occurring, as not all factors are coded immediately at the scene of the accident. The HFF model allowed for a clearer interpretation of all factors related to the collision occurring compared to the other two methods, as there were no limitations in the manner that accidents could be coded or the number of factors that could be coded. The other two models required a more rigid approach to coding cases and this presented some advantages in terms of coding similarities, but also limited the coding possibilities.

The case coding carried out in this thesis placed equal weight on variables related to the road user, vehicle and environment/infrastructure. As the statistical data mining process limited the number of variables that were entered into the analysis, the possibility of coding a large number of variables for each of the subgroups in HFF was particularly relevant. A study by Thomas, Morris, Talbot, & Fagerlind (2013) using the DREAM method on 997 crashes analysed in 6 European countries, outlined that when individual factors such as speed, alcohol or fatigue are interpreted to be the main contributor to the interaction leading to a crash, the DREAM method may not provide as much insight as other methods. This is because the coding ends after a shorter causation chain is developed compared to longer causation chains with regards to more complex cases. When the cases are more complex in nature with regards to the interaction present, the structure of the DREAM coding patterns helps provide clarity to the analysis of these cases.

The inter-rater reliability of each of the methods was quite low when compared with a group of five different crash cases provided from different countries. The DREAM method had the highest inter-rater reliability of the different coding methods, but did not have a satisfactory level of 85% or higher agreement as defined by (Krippendorff, 2004). The low inter-rater reliability ratings for the remaining two methods were considered to be due to the lack of formal training being given to the participants, other than a coding guide, and the nature of the other methods being able to code variables without linking them in the manner of DREAM. If aiming for a number of different research groups to collect accident codes, then DREAM would have been the preferable method, though as the aim was to gather as much data as possible in the data coding stage and then synthesise that data using a three-step analysis approach the HFF methods advantages were considered better in this context.

The difference in the coding approaches of the different methods led to an understanding that the analysis of these cases would be needed to be handled in a different manner for each coding method. The DREAM method allowed for the highest number of factors to be coded, but the inclusion of a causal chain approach based on previous research (Wallén Warner et al., 2008) provided an advantage in basing the codes on previous research findings, but also restricted the coding possibilities. The advantages and disadvantages of the HFF method can be considered the opposite of the DREAM method, where a large number of different variable and perceptual failure codes are possible but these coding possibilities are not based on past research. For the reasons outlined above the HFF method was ultimately selected.

3.9.5 Limitations of the research

One of the main limitations in terms of the inter-rater reliability was that only the DREAM method had an inter-rater reliability rating higher than 50% for any of the fields analysed and none of the methods had a rate higher than 70%, which demonstrates the difficulty of obtaining similar codes when coders are not provided thorough uniform training. Normally an inter-rater reliability rate of above 85% is expected when using any type of coding methodology (Krippendorff, 2004). According to Krippendorff (2004) high inter-coder agreement is reached by giving coders extensive training, allowing them to discuss their choices with other coders and adjusting problems with the coding schemes (Warner & Sandin, 2008).

It should be noted that most of the participants did not have prior training in these methods and also came from different theoretical backgrounds. These limitations may have affected the codes and cause agreement to be lower than expected.

The cases used in the inter-rater reliability and comparison studies were also collected from three different databases. Though each coding systems' aims were similar in nature, and the type of data collected was also similar, the case descriptions and coding level of detail varied and thus this may have caused coding issues to be present. Taking this issue into consideration, a comparison for case codes by different researchers rather than between the cases was carried out. The level of clarity of the accident descriptions was also varying, though for these studies the maximum possible information was collected so as to make this confounding variable not valid.

The participants were also identified as individuals that were experienced in terms of accident data collection, this was quantified as individuals that had been out to or coded at least 30 in-depth or retrospective accident cases. This could potentially limit the implications of the analysis as a more thorough understanding of the participant's level of experience and different theoretical backgrounds may be necessary to consider if there was bias in the selection made when coding cases.

3.10 Summary

In this chapter a comparison of three of the main methods of traffic accident causation analysis that are used currently throughout Europe was made. This was carried out by comparing the DREAM, ACASS and HFF accident causation methods that are used for in-depth accident investigation research.

A questionnaire study, inter-rater reliability analysis and case coding analysis was carried out using each of these three analysis methods to support the decision over the most appropriate method to be used for the subsequent analysis accident causation coding of the OTS data.

These studies helped identify the advantages and disadvantages of the different systems for accident analysis purposes. Each of the methods have been used for countermeasure development procedures in their country of origin.

According to the aims of the thesis and the results presented, the HFF method was identified as the most appropriate method for use. The reasons for the selection of HFF were identified as the HFF methods possibility to code all accident variables that can be analysed by the accident investigator and the possibility of clearer countermeasure identification.

4 The Human Functional Failure Model – A Review of the Methods and Procedures

4.1 Introduction

In this chapter the Human Functional Failure (HFF) model that was developed by IFSTTAR, France and used throughout this thesis for the analysis of accident cases with regards to accident causation will be demonstrated. A brief description of the core variables will be undertaken below. For further information if necessary please refer to two documents from the TRACE project (Naing et al., 2007; Van Elslande & Fouquet, 2007).

4.2 Human perception within causation models

In order to understand how an accident occurs it is necessary to understand the different phases that individuals go through during an accident. The perceptual stage that a road user is at before the initiation of the accident situation allows investigators to determine which possible factors caused the accident. As the driving environment is extremely complex, the road user can only perceive a limited amount of information from the environment. The driver in turn needs to select the most relevant information from the traffic environment in order to make the necessary driving actions. After selecting the necessary information then an interpretation of the information needs to be made in order for a decision process to be made. These decisions are based on previous knowledge of different situations (Van Elslande & Fouquet, 2007). The driver then needs to carry out the behaviours that are necessary for the action to be undertaken. Thus this is the point that functional failures can occur. This is termed 'human error' in the ergonomics field (Reason, 1990; Van Elslande & Fouquet, 2007).

When an error or violation occurs before an accident, it depends on which stage of perception the road user is at. Within the traffic environment the road user is constantly interacting with the other road users and constantly reacting to what is happening around them (Van Elslande & Fouquet, 2007). For each specific situation a decision needs to be made, and a breakdown or incorrect decision can lead to either a near miss or the occurrence of an accident. Feedback loops are also necessary to be considered, when identifying traffic flow patterns, as a way of using countermeasures for drivers in situations where there decision making process needs to be altered for safety precautions.

Van Elslande and Fouquet (2007) defined a five stage perceptual model related to how road users perceive their environment. The road user is constantly interacting with the environment around them, and going through a loop of five stages when perceiving traffic. The road user first perceives (stage 1) the information from the environment, then diagnoses (stage 2) the situation, anticipates (stage 3) how events will unfold, makes a decision (stage 4), and then performs an action (stage 5). So in essence we can identify the 5 stages in a progressive manner which can be seen in figure 11. When considering these groups it is beneficial to understand them in terms of Norman (1981), Reason (1990) and Rasmussen's (1982) work on human error. Violations both made deliberately and as a result of the situation can be situated in the decision stage (4). The effective understanding of the dynamic traffic environment can be seen in two stages in the diagnosis (stage 2) and the prognosis (stage 3) stages, where a lapse in awareness can result in error or behaviours that will lead to the triggering of an accident. This can be viewed in terms of Endsley's (1995) work on situational awareness, as the diagnosis is related to a situational understanding of the traffic system and prognosis the expectations that the road user has and how 'aware' they are of what is going on around them with regards to interactions. The other stages are related to slips and errors.

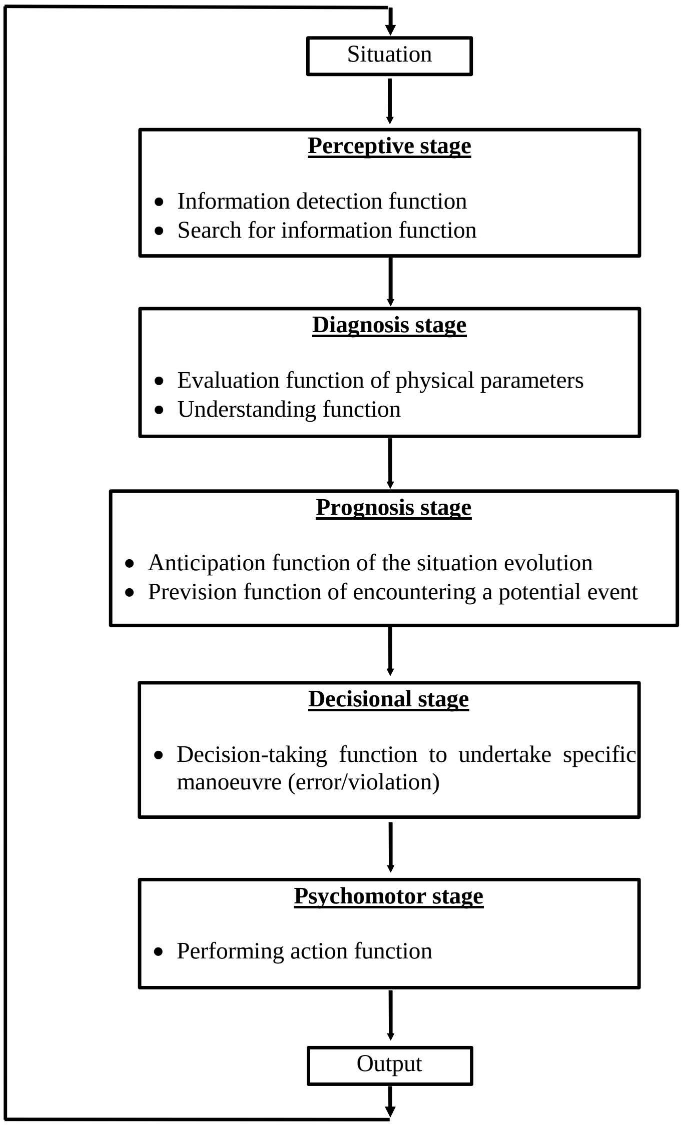


Figure 11: Functional chain involved in driving activity (Adapted from Van Elslande and Fouquet, 2007)

In the section below a description of the 20 sub-groups of human failure can be seen. All of these descriptions and the following descriptions with regards to human, vehicular and infrastructure factors are taken directly from Van Elslande and Fouquet (2007) and Naing et al. (2007). All of the perceptual stages and specific failure types coded can be seen in figure 12.

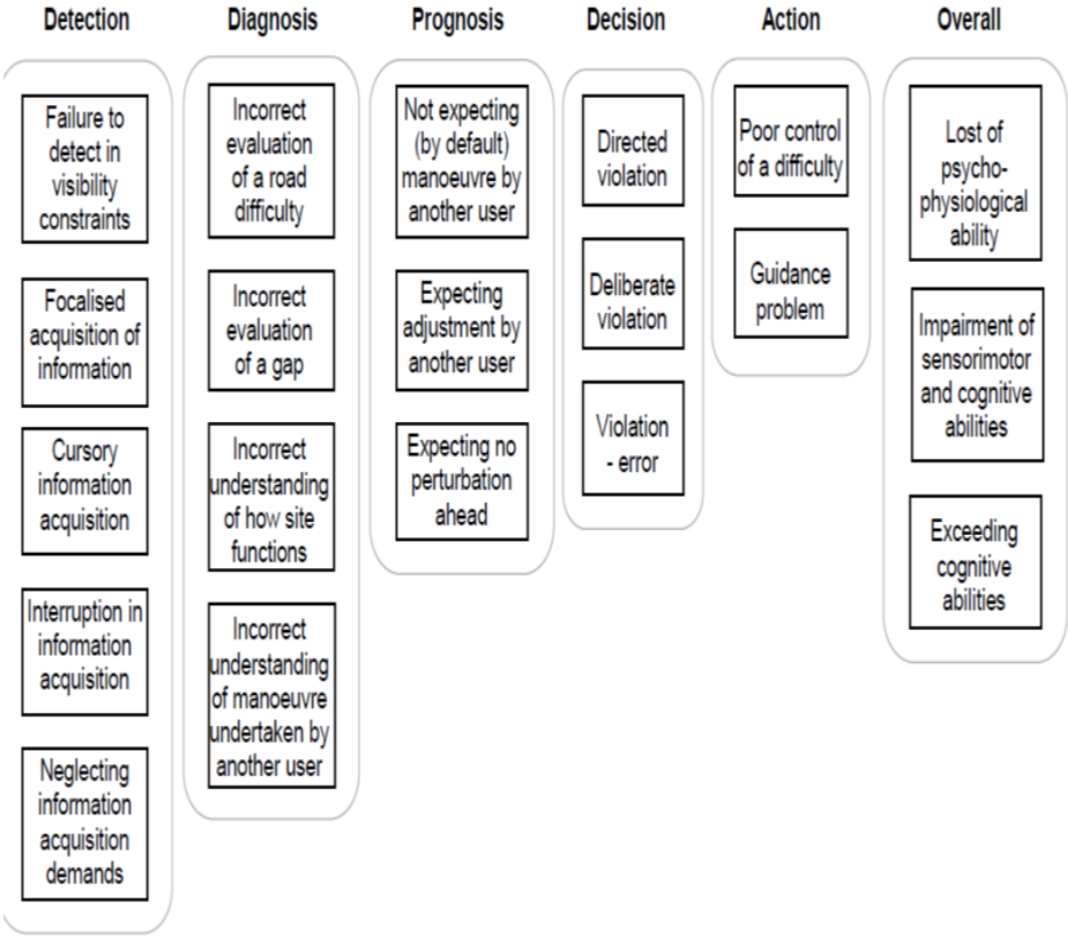


Figure 12: Perceptual stages and failure types for road user (Adapted from Van Elslande & Fouquet, 2007)

Detection failures

1. Non-detection in conditions with limited visibility: This failure is due to either an environmental or vehicular constraint limiting the driver's ability to detect an important object/situation during driving (e.g. as a result of night or the vehicle infrastructure effecting visibility).

2. Focusing on a part of the visual environment causing incomplete information acquisition: This type of failure occurs when a road user focuses their attention on a particular or complex problem (without a conscious choice) during the journey and so does not detect an

object that needed to be detected (e.g. failing to detect a moving vehicle).

3. Cursory or hurried information acquisition: This failure occurs when insufficient time is given to the visual field resulting in a failure to detect a hazard. Factors could include a busy traffic environment or a rapidly changing traffic situation.

4. Momentary interruption in information acquisition activity: This failure occurs as a result of distraction from inside or outside the car but not related to the driving task, for example the monotonous nature of the driving task resulting in a loss of attention.

5. Neglecting the need to search for information: This failure is due to the road user not searching for information when it was required as they believe that it is not necessary, for example when the driver has the right of way or is familiar with the road.

Diagnosis failures

1. Error in evaluating a passing road difficulty: These are failures that occur as a result of incorrectly evaluating a road difficulty.

2. Error in evaluating of the size of a gap: These are failures that occur when identifying the speed and distance of a vehicle that is travelling in the same direction or direction other than the road user.

3. Mistaken understanding of how a site functions: These are failures that occur as a result of not understanding the signs/layout of an area either due to the road user or the layout

4. Mistaken understanding of another user's manoeuvre: These are failures that occur as a result of the other road user giving misleading or ambiguous information or the main road user only interpreting the other road user’s behaviour briefly and thus incorrectly.

Prognosis failures

1. Expecting another user not to perform a manoeuvre: In the absence of cues to the contrary, drivers who have the priority at an intersection do not expect a non-priority user who is stationary to start moving forward and are surprised by this unexpected movement.

2. Actively expecting another user to take regulating action: The driver expects another user to either stop or undertake a precautionary behaviour, though the other driver does not do this.

3. Not expecting to encounter another road user ahead: The driver adopts a behaviour that does not integrate the possibility of encountering an impediment, despite a lack of visibility.

Decision failures

1. Violation directed by the characteristics of the situation: The driver is confronted with a situation in which they were directed to take a certain level of risk in order to attain their goals.

2. Deliberate violation of a safety rule: This driver performs a behaviour that can be deemed as 'risk-taking' in the performance of a manoeuvre.

3. Unintentional violation: When a road user makes a behaviour that is unintended but is still a violation.

Execution failures

1. Poor control of an external disruption: In situations where s/he meets severe constraints, drivers are no longer able to control the trajectory of their vehicle. These failures arise either from an external disturbance (skid on wet or icy road, presence of a wasp in the passenger compartment), or from a sudden mechanical defect (defective brakes, breaking of a cable which comes to be stuck in the front wheel of a motorcycle, etc.).

2. Guidance problem: The road user undertakes a secondary activity (picking up fallen objects inside the vehicle, talking to a passenger in the rear, and so on), and veers off course.

Overall failures

1. Loss of psycho-physiological capacities: A loss of awareness by the driver as a result of being taken ill or falling asleep because of tiredness and/or a high blood-alcohol level, which occurs most frequently during a simple guidance task.

2. Alteration of sensorimotor and cognitive capacities: Even if they did not fall asleep as above, the drivers showing this capacity failure are not in a psychophysiological state for adequately controlling driving activity as a whole.

3. Overstretching cognitive capacities: Linked to a general lack of skill in relation to driving activity (age, occasional driving), drivers find their abilities are overstretched the moment they encounter a difficulty in their progress, and they sometimes carry out absurd manoeuvres.

4.2.1 Pre-accident driving situations

This is the first part of the HFF analysis. The method identifies certain driving manoeuvres to classify the driving situation immediately before the crash. The list of behaviours is as below:

1. Going ahead: The user was 'going ahead' and not making any specific manoeuvres prior to the rupture phase.

2. Changing lane: The user changed lanes into another lane travelling in the same direction, but was not overtaking another vehicle.

3. Overtaking: The user was overtaking another road user or a stationary vehicle.

4. Turning: The user was making a turning manoeuvre (e.g. at an intersection, U-turn).

5. Stopping: The user was stopping/slowing in the carriageway (e.g. parking, approaching stationary traffic queue, approaching a junction/traffic control, slowing to turn into driveway/side road).

6. Reversing: The user was reversing (e.g. on main carriageway, into side road/private drive, out of side road/private drive, into roadside parking space).

7. Starting: The user was pulling away from a parking space/driveway/junction/traffic control/traffic queue.

From investigating the different types of accident locations defined in accident data collection systems across Europe and Australia a number of general location types were identified. The manoeuvres listed above were put into the relevant behaviour list so as to differentiate between each specific manoeuvre that a road user made. This list is as below:

1. Going ahead
    a. Straight road
    b. Road with bend
2. At intersection:
    a. Give way
    b. Stop
    c. Traffic signal controls
3. Manoeuvres
    a. Overtaking
    b. Lane changing
    c. Slowing
4. Other locations
5. Roundabout

6. Slip road

7. Pedestrian crossing (not at intersection)

8. Railway crossing

Along with the manoeuvre undertaken by the road user, these locations form part of the driving situation.

4.2.2 Contributory factors related to the human

As the main failure describes the specific failure that led to an accident occurring, it is also necessary to identify different type of factors that contributed to the failure developing and ultimately occurring. Below is a list of the contributory factors that cause accidents to occur.

A. User state

The ‘state’ of the user includes physical, physiological or psychological conditions, either pre-existing or brought on by substances taken, such as alcohol or drugs.

A1. Physical/Physiological

The physical or physiological state of the user can have a major effect on the outcome of a potential accident situation. Often, the danger signal is never perceived, because either the road user does not know they have a medical condition or the user does not realise that their pre-existing ‘state’ puts them in a position of having a higher likelihood of a failure occurring.

A2. Psycho-physiological condition

The psycho-psychological condition of the user will also have a major influence over the potential for functional failures to occur, as will any substances they have taken. These factors include any substances taken or whether the road user was emotional, fatigued or in a hurry.

A3. Internal conditioning of performed task

These factors are related to the task that the driver is performing, but refers more specifically to the ‘conditioning’ of the driver to the task (e.g. the informal rules the driver follows, either consciously or sub-consciously).

A.4 Risk taking

These behaviours are intentionally risk taking. The road user is normally aware of the ‘chance’ they are taking but, for other reasons (e.g. experience, substances taken) they still choose to proceed with the action. Types of risk taking include speeding, vehicle positioning, following distance of other vehicles (e.g. time headway the distance or time between vehicles), traffic control being disobeyed and eccentric motives.

B. Experience

The user’s prior exposure to the task at hand or their surroundings will affect the way they process information. The factors here are whether the road user had too little or too much experience either driving for the particular roadway/type, for example driving on the left-hand side or driving in a new area.

C. Distraction

The behaviour of the road user can affect the way they control their vehicle and respond to both their internal and external surroundings. Three types of distraction factors are identified: (1) distraction within the user, (2) distraction outside the vehicle, and (3) distraction inside the vehicle.

4.2.3 Contributory factors related to the environment and Infrastructure

This factor encompasses all aspects related to the users’ surroundings (e.g. external to the vehicle and road user). Six categories of environment related factors have been defined and are outlined below:

A. Road condition

The condition of the road surface will affect the road user’s ability to be able to control their vehicle on the road. The condition of the road will be affected by the contaminants and defects, plus the road surface type itself.

B. Road geometry

The layout of the road itself will also affect the road user's ability to control their vehicle.

C. Traffic condition

The flow, speed or density of the traffic on the road will potentially affect the road user's ability to undertake their journey.

D. Visibility impaired

If the road users visibility of the road ahead is impaired in some way, this will undoubtedly increase the possibility of a functional failure occurring.

E. Traffic guidance

If there is a fault or a failure in the traffic guidance system (signs, traffic signals and road markings, including reflective studs and painted lines), this will affect the road users ability to undertake the driving task.

F. Other environmental factors

Obstacles and other factors which suddenly appear within the road/roadside will affect the road user's ability to undertake their journey, even when an impact does not occur with these obstacles.

4.2.4 Contributory factors related to the vehicle

The vehicle factors encompass all aspects related to the vehicle used. Possible factors influencing the vehicle include:

A. Electro-mechanical

Electro-mechanical factors are 'failures' which directly affect the vehicle's control. This type of failure would generally result in it being physically difficult/impossible to control the vehicle.

B. Maintenance

Maintenance factors are anticipated vehicle faults, indirectly affecting the control of the vehicle. They may make it more difficult (e.g. in terms of visibility) or 'illegal' to drive/ride the vehicle, but it is still possible.

C. Design

Design factors are those related to the ergonomic design of vehicle, which affect its safe/efficient operation by the road user.

D. Load

These factors relate to the load of a vehicle involved in an accident. If a vehicle drove into another vehicle ‘poorly secured’ discarded load, this would be an ‘obstacle in the road’.

4.2.5 Degree of involvement of the driver

This variable defines the role played by the driver in the formation of the accident. Close to the notion of 'responsibility', it differs from this by the reference not to a legal code but by the recourse to a strictly behavioural reference. In an ergonomic approach, we try only to clarify the respective degree of participation of the various users involved in the same accident. Three separate possible selection groups are defined.

Primary active

These groups of drivers are those that are primarily causing the triggering of the episode. They are directly at the origin of the destabilization of the situation. Following the functional failure, the drivers initiate for themselves or for the other involved users in the system, a critical situation in which the accident situation is going to take place.

Secondary active

These drivers are not at the origin of the disturbance which precipitates the conflict, but they are part of the development of the accident situation, by not trying to resolve this conflict. They are not attributed a direct functional implication in the destabilization of the situation, but they were a part in the non-resolution of the problem by a wrong anticipation of the events evolution.

Non-active

These drivers are confronted with an atypical manoeuvre of others that is hardly predictable, whether it is or not in contradiction with the legislation.

They are not considered as 'active' subjects because the information they had did not enable them to prevent the failure of others. They were not able to anticipate, for lack of information, the development of the situation, while the avoidance of the accident would have been possible in theory if this information had been supplied to them in time.

4.2.6 Accident configuration types

The method used identified different accident configurations in a diagrammatic form in order to demonstrate the failure sequence in its interaction stage. The coding method was developed in France by the Laboratoire d'Accidentologie (LAB). The coding document can be found in Appendix C (PP383-387). This coding method identified 8 specific grouping types:

1. Accidents with vehicles driving in the same direction
2. Overtaking accidents
3. Accidents at intersections (Including roundabouts and merging roads)
4. Accidents occurring with a vehicle leaving a parking space
5. Single vehicle accidents
6. Main accident occurring after a previous impact
7. Special cases
8. Pedestrian accidents

A large number of scenarios were available to select for each of these specific accident types. After the scenarios identified by the descriptive analysis carried out within each of the analysis chapters accidents that had similarities were grouped together in ways to be meaningfully interpreted, as the large number of diagrams would not yield significant results otherwise.

4.3 Summary

The purpose of chapter 4 was to present the analysis processes that were undertaken in this study (HFF and LAB coding). The coding structure and use of the coding sheets were explained and the different possible codes were identified. This system was used for all case coding in chapter 6, 8 and 9. Each accident was coded using the full coding sheets but only factors deemed relevant for each study were included in the analysed data.

5 Methodology

This chapter will discuss the methodology that will be applied in this thesis. It will be divided into four sections discussing the participants, instrumentation/ measures, procedures and the statistical analysis that will be used throughout the following chapters. The research design that was used in this study was a cluster analysis design that aimed to identify similarities between data obtained in a real world setting.

5.1 Participants

5.1.1 On the Spot study

The UK On the Spot (OTS) study was carried out between the years 2000 – 2010 with the aim of collecting in-depth accident data on the scene of an accident. This project was carried out aiming to continue in the essence of on the spot studies carried out in the UK by Starks and Miller in 1961 at the Road Research Laboratory (RRL), Mackay in 1964 who formed a multi-disciplinary team working closely with the Birmingham Accident Hospital (Mackay et al., 1960) and the TRRL studies carried out between 1970 and 1974 in the area in and around the Transport and Road Research Laboratory in South East Berkshire, UK (Sabey & Staughton, 1975).

For the OTS study accident data was collected by two separate groups in different areas within the UK. The Vehicle Safety Research Institute (VSRC) collected cases within the South Nottinghamshire area of East Midlands, England and the Transport Research Laboratory (TRL) covered the Slough, Reading, Henley on Thames and High Wycombe areas in the South East of England.

5.2 Procedure

5.2.1 On the Spot study and data gathering

Within the OTS study accident researchers responded to calls four times a week during 8 hour shifts, alternating hours for random data, to accidents that happened within this area (Morris et al., 2006). The study collected a total of 4,004 accidents involving 12,749 vehicles and 527 pedestrians. The OTS team consisted of at least 2 accident researchers and a police officer. Accident researchers reported all relevant data in terms of the vehicle, environment, infrastructure and human participant in relation to the accident. They also deduced and reported causal factors that were related to the formation of the accident process.

Expert research teams attended the scene of road accidents, typically within 20 minutes of the incident occurring to make an in-depth investigation that included the highway, vehicles and human factors that were present. Data was also collected retrospectively after the accident occurred. The first step once on scene required the serving police officer on the OTS team to make contact with the police officer in charge of the accident scene, explaining the OTS procedure and intended activities that will be carried out. After the fulfilment of protocols and safety requirements, the team made contact with the people and the various elements involved in the crash. Data was coded in a library of some 200 forms with over 3,000 individual variables (Gkikas, 2009).

Photographic evidence of the accident scene as well as physical measurements were also recorded. If possible the researchers on scene had a short interview with the accident participants which were not recorded electronically. Witness statements were also gathered where possible. A questionnaire was also sent out to applicable road user groups asking questions relating to background information and the accident description, the response rate of these questionnaires was close to 50%.

The OTS study concentrated on gathering information with relation to human factors, vehicles and infrastructure. The OTS study aimed at acquiring volatile data from the crash scene (Hill & Cuerden, 2005). The investigators first gathered data from vulnerable road users, where applicable, and then collected volatile evidence on the highway such as contact marks, trace marks and damage to road features. The vehicles were examined, with smaller vehicles being examined first, and finally measurements of the environment were taken and video and photographic recordings of the accident scene were made (Hill & Cuerden, 2005).

This was carried out by the accident researchers for the accident on scene. For the vehicle, as much information with relation to the vehicle were recorded as possible on scene in terms of the vehicle, any marks or changes on the vehicle body, the tyres and all other objects on the vehicle.

Relevant information was also gathered in terms of the highway and information from the area where the accident occurred and 50 metres before and after that area were collected, with the aim of data wholeness. The human data was collected through observations and where possible interviews with the individuals involved in the accident. This information was then put into the relevant human factors and accident causation fields in the database.

Accident reconstruction information was also carried out and put into the dataset. This information was produced through interpretation of the accident data and where possible (and applicable) a PC crash simulation was undertaken.

The sampling procedure used was based on stratified random sampling and made up the sample according to sponsor recommendations (Morris et al., 2006). The two teams remained on standby for a nine-hour shift period ready to respond immediately to an accident notification. The shifts were devised as a rotating system to ensure that the dataset gathered could be statistically weighted with regards to national data (Cuerden, Pittman, Dodson, & Hill, 2008). Shift patterns consisted of six days on and four days off.

5.2.2 On the Spot (OTS) causation measures method

For the On the Spot team the police driver filled out the forms in relation to the path the driver took and the collision codes, while the accident investigators filled out the interactions codes and injury codes on a concurrent level. The OTS study aimed to identify contributory factors in a road accident and the key actions and failures that lead directly to the actual impact. The causation coding systems used within OTS were largely subjective and depended on the skill and experience of the investigator to reconstruct the events which led directly to the accident (Hill & Cuerden, 2005).

OTS used 5 different coding systems for describing accident causation:

1. The 1995 UK police system: aims at determining (1) the critical failure of the manoeuvre and (2) factor(s) that caused this failure.

2. Causative features: the investigator selects a feature of the accident and selects whether it was (1) definitely, (2) probably ,(3) possibly, or (4) not causative. There can only be one coding per vehicle for each factor.

3. Crash causation code: 20 variables that explain why a crash happens in relation to the driver are able to be selected.

4. Interaction codes: Are divided into 7 categories (1) legal; whether the driver did anything illegal (e.g. above speed limit, not obey a sign or was above legal alcohol limit (2) perception; related to what the driver was expecting, looking or planning for that driving situation (3) judgement; after the perception level what the driver decided and how he/she acted according to the driving event. (4) loss of vehicle-control; loss of control due to several different factors (braking, acceleration, cornering etc.) (5) conflict; explains the conflict level of the accident, when the accident situation arose (6) attention; describes if there were any distractions or general inattentiveness (7) impairment; describes performance impairment due to illness, substance abuse, fatigue or other factors.

5. Self-reported assessments (questionnaire): allows individuals involved in the accident to state what they believed caused the

accident in a self-report method (Lenard & Hill, 2004). These reports are used as additional information with regards to the crash.

5.3 Measures

5.3.1 HFF accident coding analysis

For the research conducted in this thesis all of the cases that are included in the OTS analysis chapters were coded retrospectively by the author using both the HFF and LAB accident type coding analysis methods. All cases were coded by the author by looking at all of the relevant data within the OTS dataset, which included:

- Detailed descriptions of the cases by the crash investigators
- Case notes that provided detailed information on each of the cases, including interviews and witness statements that were gathered
- Detailed measures for all possible factors with regards to the road user
- Detailed measures for all possible factors with regards to the vehicle
- Detailed measures for all possible factors with regards to the environment and road infrastructure
- Accident causation coding using the OTS causation measures as described in section 5.2.2

The coding of each case took a minimum of 30 minutes, and a total of 2086 cases were coded. The OTS coding provided a traditional expert coding of cases and the HFF coding further elaborated on this by coding all relevant factors from the accident. An analysis of all powered two wheeler accident cases (chapter 8) and pedestrian accident cases (chapter 9) present in the OTS dataset between the years 2000 to 2010 was carried out using the full HFF coding model and LAB accident coding type system for this thesis. For the all accidents analysis (chapter 6), all cases collected by the OTS team between the years 2000 to 2003 was coded using the full HFF coding method and LAB accident coding type system. The OTS causation coding system uses a 4 point system coding the causes as definite, probable and possibly contributory. The accident was analysed in terms of factors that were deemed to have been a factor in the causation of the accident, only the factors that were identified as definitely influencing the accident were included in the analysis regardless of the original coding by the OTS accident investigators.

The reason that the years the cases coded were different for the different analysis chapters was the numerical requirements of the chosen statistical methods. For each dimension of analysis within a cluster analysis 10 cases need to be included. If 10 dimensions were included 100 cases would be needed to be provided, if 100 dimensions are included 1,000 cases would need to be included. As the number of clusters increases the number of cases needed would also increase. A minimum number of cases were needed to be met for analysis purposes and as the PTW and pedestrian accident cases were smaller in number an analysis of all possible cases was needed to attain statistical significance.

These analysed cases were then merged with the relevant OTS files in order to use all of the acquired data where necessary in the analysis and understanding of these accident types. The coding sheets used for the HFF and LAB coding procedure can be found in Appendix C (PP362-387).

5.4 Statistical analysis

5.4.1 False starts

During the span of study a number of different statistical procedures were considered and preliminary analysis with the data was carried out using these procedures. Two statistical procedures were attempted to be used with the sample data. The procedures that were attempted to be used were:

1. Principal component analysis

2. Quasi induced exposure methodology

Principal component analysis (PCA) is a statistical methodology that is used to reduce large and complicated multivariate datasets into a simpler form. This analysis allows multiple factors to be linked to a set of components that explain the variance in the group of cases, allowing for correlations between these factors to be analysed and interpreted.

Rather than using pre-conceived data chains to analyse the data this allows for an exploratory analysis of the factors linking them to particular detection stage failures. This type of analysis is most commonly used with questionnaire data, in order to group the questions into specific groupings. The aim of PCA is to analyse the variance of the factors that adds up to 80% though with accident data usually 50% is seen as a significant number for analysis. Two main limitations are present for PCA, the first is that it limits the number of variables that can be entered into the analysis. The second is that it does not provide groupings that will help identify plausible large groupings for scenario building.

An attempt was made to use PCA with the OTS accident data but the output only provided at most two categories/components that were interpretable with regards to scenarios and the rest of the categories were moved to other axis. This did not allow for a clear interpretation of the data as the largest factors were included in two axes and from the third axis onwards it was not possible to find any meaningful results, which resulted in a loss of information.

The quasi induced exposure method uses road users deemed to be ‘not at fault’ as the comparison group to road users that are identified as entirely at fault. This method selects only two-vehicle crashes in which one driver is declared entirely responsible and the other entirely not-responsible compared to the case control method where all crash types are included (Lenguerrand, Martin, & Laumon, 2006). Drivers that are not assigned any human factor causes within a crash situation are identified by the investigator as being ‘not at fault’. This method assumes that the not at fault drivers represents all drivers exposed to the crash hazard, and so represents the total population. This allows for the two separate groups that are being compared to have the same comparable variables and thus allows a more comprehensive analysis of accident data as a single factor methodology (Chandraratna & Stamatiadis, 2009).

Chandraratna and Stamatiadis (2009) identified two points of limitations in regards to the quasi induced methodology, firstly that not at fault drivers may be coded incorrectly and in fact are either partly or fully at fault in an accident. In terms of accident research we assume that an accident researcher correctly interprets an accident scene as a result of both background and study, so this would be a limitation of most accident research. Secondly this method cannot explain single vehicle accidents as in these types of accidents the road user is usually at least partly to blame for the occurrence of the accident.

Furthermore in the analysis of accident causation in traffic accidents it is necessary to not administer fault to any of the involved road users, as this would not allow for solutions to be described for road users that would be deemed as ‘not at fault’ in the quasi induced method. These methods found to only provide numerical values in terms of risk rather than a complete scenario setting, and so were not used for this reason. Due to the nature of their analysis requirements they could also not be used with single vehicle accident settings, which was an analysis that was also carried out in this thesis.

5.4.2 Latent class cluster analysis

Cluster analysis methods can be considered as data mining tools which are placed between statistical methods and information processing. The main aim of cluster analysis is to differentiate objects in groups by identifying similar objects and putting them in the same group depending on the variables entered into the analysis (Rezankova, 2009). The similarity is dependent on quantitative or qualitative variables and how similar each group is depending on these features (Rezankova, 2009).

Cluster analysis is most commonly used to maximise the similarities between in cluster elements and the differences between inter cluster elements

(Fraley & Raftery, 2002). These measures are called similarity based clustering methods and use a distance function for continuous variables. In the situation that variables consist of continuous or qualitative elements then the variable can be mapped onto a binary measure to enable comparison (Depaire et al., 2008).

Among traditional similarity based clustering approaches two major approaches can be seen, the hierarchical approach (e.g. Ward's method, single linkage method) and the partitional approach (e.g. K-means) (Depaire et al., 2008). Hierarchical cluster analysis types use a distance measure to be able to handle data. When the data is mixed in nature (includes both categorical and continuous data) these measures may not yield satisfactory results as the analysis will try to analyse categorical data by comparing distances and the distance functions in the continuous data may not be applicable for the categorical data, which will in turn alter the results. K means clustering works on the basis that each cluster is partitioned into a number of clusters represented by their centres (or means). The mean of this analysis needs to be defined and it is particularly sensitive to noisy data and outliers (Rokach & Maimon, 2005).

The principle behind each cluster method is similar in that all cases are firstly considered as individual clusters. Clusters are merged depending on the type of method chosen, as each method has a specific criterion to merge cases. In all methods we begin with as many clusters as there are cases and end up with just one cluster containing all cases. By using mathematical procedures it is possible to identify clusters that are highly similar to each other (Field, 2013). The data used in this clustering can only be interval or binary data. The reason for this is that the clustering method uses a distancing measure to group cases together and so treats cases that are categorical in an interval nature. As most of the data collected from traffic accidents is categorical these methods were not viable.

The second type of clustering methods are known as latent class modelling. In this model every cluster has an underlying probability distribution to generate data (Depaire et al., 2008). These models can be recast as a statistical choice procedure, and a comparison of different models is possible with goodness of fit measures (Fraley & Raftery, 2002).

Latent class analysis is a statistical technique that can be used to analyse multivariate categorical data. When data has a number of categorical variables it is often of interest to identify cases that contain similar aspects, and identify whether these similarities hold over different variables (Linzer, 2008).

Latent class models can be used to undertake the above stated goals. According to Linzer (2008) "The latent class model seeks to stratify the cross-classification table of observed variables by an unobserved unordered categorical variable that eliminates all confounding between the manifest variables. Conditional upon values of this latent variable, responses to all of the manifest variables are assumed to be statistically independent, an assumption typically referred to as "conditional" or "local" independence" (p. 1). The variables that are selected to be included in the model should not be completely identical to other variables and have a measurable difference (Linzer, 2008). In the case that a model includes outliers an extra cluster is added to include these cases and separate them from the model (Fraley & Raftery, 2002). Typically this group of outliers can be identified if a cluster is hard to profile by means of the distribution within the analysis (Depaire et al., 2008).

The model, in effect, uses probabilistic grouping to produce expectations with regards to how the observation will respond on each manifest variable. Although the model does not automatically determine the number of latent classes it offers a variety of parsimony and goodness of fit statistics that can be used to make a theoretically and empirically sound assessment (Linzer, 2008).

Some of the limitations for this cluster analysis method are:

1. A good understanding of the data is necessary as only relevant variables are needed to be included in the analysis for the results to be meaningful.

2. An interpretation of the significance of the results is needed to find meaningful relationships.

3. Large sets of data are needed so that the analysis remains significant and meaningful.

5.4.3 Chi square test

After the cluster analysis was carried out each cluster was compared to the total values for each individual variable using a chi square goodness of fit test. This test is used to test whether observed data follow a particular distribution. The chi square statistic for this analysis is demonstrated in figure 13.

$$X^2 = \sum \frac{(\text{observed - expected})^2}{\text{expected}}$$

Figure 13: Chi Square goodness of fit test

The results of the cluster analysis were analysed with a Pearson's chi square goodness of fit test for each of the factors included in the analysis and the specific values of these factors. The observed counts for the variable and the expected counts for this variable are compared with the observed counts for the same variable against the same dataset and the formula demonstrated in figure 13 is used to see if there are any differences for the values. A chi square goodness of fit test is carried out to see if there is significant difference between the groups within the data and an analysis of the value residuals is carried out to determine where the difference is found.

The chi square test was carried out by comparing the percentage in the cluster for the value against the overall value and then calculating the residual variance. Values that were over 2, meaning that a factor was over-represented, were included as significant results for all of the factors (and their values) that were included in the cluster analysis.

For example if we take gender as the variable, the values for male and female in cluster 1 are compared against the dataset as a whole identifying the observed and expected values. A significance test is then carried out to see whether there is a significant difference within the cluster compared to the overall values for the variable. If there is a significant difference a residual test to see where the difference is from determines whether the positive difference is coming from males in the cluster or females and whichever had a value over 2 from the residual test is taken as an over-represented value. This procedure is then carried out for all variables within the cluster model.

This test allowed for it to be determined whether the value of a variable was over-represented or under-represented. For example if one of the variables entered into the cluster analysis is gender and for the specific cluster females attribute for 40% of the individuals while for the full dataset they account for 20% the chi square analysis identified whether being female was significantly over-represented for the specific cluster. This test was carried out for all clusters and all variables in the dataset.

Though these significant results were taken into account for the analysis the overall frequency values were also kept in mind when analysing the results, as a factor being over 75% in average would require a value of close to 80% depending on the number of cases to be considered significant, though in essence would still be over-represented when considering the descriptive values separately.

5.5 Statistics software

The statistical analysis for this thesis was carried out using IBM SPSS Statistics version 21®. The OTS data was made available by the Transport Safety Research Centre (TSRC) in Loughborough University. The HFF coding sheets were coded in excel, then converted into SPSS form and merged with the OTS database by the author.

The LCC analysis was conducted using the R statistics program which is an open source programming language and software environment. SPSS® and

R are able to be integrated when the related packages are used. R version 2.14 is the applicable program for SPSS version 21 and so these two statistical analysis programs were integrated using the programs available at the IBM website. Due to the nature of cluster analysis including groups with values that were too small would cause the cluster analysis to give results that were not meaningful, thus an analysis of the descriptive data was carried out for each of the analysis chapters separately and values of the variables were grouped in meaningful ways both based on the literature review and the HFF coding categories provided.

A package that would be able to carry out Latent Class Cluster analysis was identified for the R statistics program as poLCA, which was developed by Linzer (2008). The poLCA package allows an analysis of both latent class clustering and latent class regression methods.

5.5.1 Clustering algorithm

According to Eshghi, Haughton, Legrand, Skaletsky, and Woolford (2011) the basic latent class cluster algorithm can be given as demonstrated in figure 14;

$$P(y_n|\theta) = \sum_{1}^{s} \pi_i P_j(y_n|\theta_j),$$

Figure 14: Latent class cluster algorithm

Source: Ehsghi et al. (2011)

Ehsghi et al. (2011) defined this formula as “Where y_n is the nth observation of the manifest variables, S is the number of clusters and $\_j$ is the prior probability of membership in cluster *j*. P_j is the cluster specific probability of y_n given the cluster specific parameters Θ_j. The P*j* will be probability mass functions when the manifest variables are discrete and density functions when the manifest variables are continuous” (p. 273) (Figure 14).

5.5.2 Analysis process

The analysis process used in the analysis chapters using the OTS data consisted of the same systematic process. First each individual OTS case was coded using the HFF and LAB coding sheets by analysing the OTS coding as source material. Once all coding was completed the new coding sheets were merged with the SPSS files from the OTS data and a latent class cluster analysis using the poLCA package was carried out using SPSS. This package was run through SPSS using syntax (which can be seen in Appendix A, Page 328) first to compare 2 to 15 cluster solutions with regards to the AIC and BIC measures, and then to identify the cluster that was statistically relevant in terms of goodness of fit. This cluster was then printed out onto an excel sheet and a detailed chi square analysis measure was carried out for each of the clusters identified versus the total values for the data in order to identify significant values for the factors within the clusters. These significant clusters were calculated based on chi-square significant levels of 95% confidence intervals ($p<0.05$). These results were then interpreted using a systems approach to identify applicable countermeasure indications for the different failure scenarios described.

Certain statistical measures were undertaken in order for the LCC analysis to be interpreted. A descriptive analysis was conducted for each specific accident set that was studied. From this analysis factors that were identified as having a high frequency were included into the data where applicable.

Latent class clustering (LCC) methods requires the user to manually define the number of classes for the analysis sample. The number of clusters ultimately selected is related to certain statistical measures that are available in the data that describe the goodness of fit of the model.

When selecting this model, a Bayesian model evidence is carried out. A statistical model identifies how one or more factors are related to each other and the goodness of fit (GOF) describes how well a model fits a set of observations and summarizes how different the observed and expected values within the model are.

The selection of the cluster analysis was carried out using two goodness of fit measures, (1) the Bayesian information criterion (BIC) and (2) the Akaike information criterion (AIC). AIC is an estimate of a constant plus the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model, so that a lower AIC means a model is considered to be closer to the truth (Wang, December 4, 2014).

BIC is an estimate of a function of the posterior probability of a model being true, under a different Bayesian setup, so that a lower BIC means that a model is considered to be more likely to be the true model (Wang, December 4, 2014). Through the nature of BIC it has more of a chance of choosing too small a model while the opposite is true for the AIC. In cases where the BIC and AIC are different the BIC would be smaller and the AIC should be larger when considered. The lower the numerical value of these procedures the better the fit of the model to the data.

When selecting an appropriate model from the two, research carried out by Lin and Dayton (1997) identified the AIC as more appropriate than the BIC when there are complex models that include a variety of different factors and groupings, of the type that is encountered in this research. The AIC should also be preferred unless there are more than several thousand cases or the sample is based on a few criteria (e.g. variables), in which case the BIC is preferred (Lin & Dayton, 1997).

The LCC method also provided a residual degrees of freedom measure that is based on the number of observations of the factors in the cluster against the number of cases used in the analysis. In instances where the number of observations is greater than the number of cases, this will cause issues in the validation of the model. Thus the number of cases and factors chosen were selected so that cluster analysis between 2 and 15 clusters could be interpreted and compared with regards to the AIC and BIC.

This thesis aims to find patterns and fully understand accident data, so it was more beneficial for there to be too many clusters than too few thus the AIC and BIC were both observed and the AIC was chosen as the definitive measure for model selection, especially considering that real world accident data collected with a sampling strategy was used in this analysis. The complete process of analysis can be seen in figure 15, which demonstrates the step by step approach that was taken in order to analyse the accident data and carry out the latent class cluster analysis,

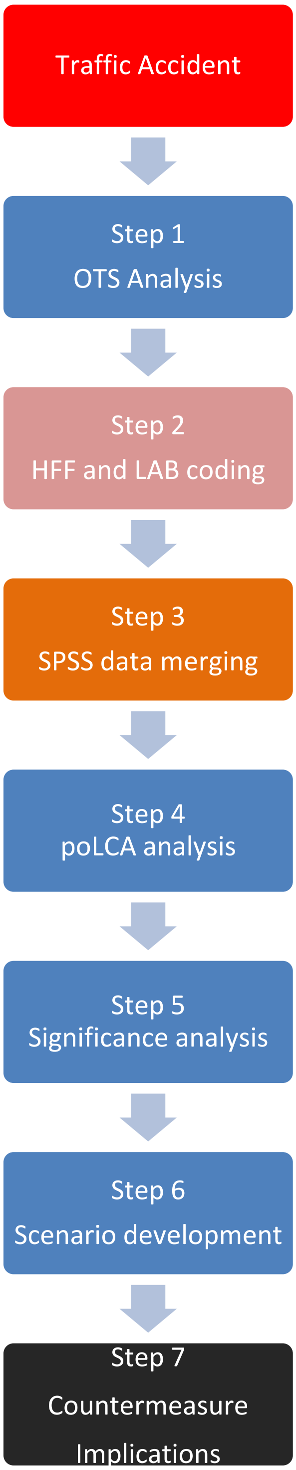


Figure 15: Analysis process of cases for the thesis

5.6 Summary

The aim of chapter 5 was to describe the methodology that was used in this thesis describing both the data gathering process and statistical analysis hybrid procedure that was put in place to help identify accident scenarios in a detailed manner. The data collection procedures that were used in the OTS study were described in some detail, identifying how accident researchers gathered relevant information on each particular aspect of a traffic accident.

Each case was then analysed retrospectively using the HFF method to gather relevant accident causation information for each accident. This data was combined into an SPSS data file that would allow a latent class cluster analysis to be carried out and reported.

The latent class cluster analysis was run using the poLCA package within the R programming language using syntax to run the analysis through SPSS. The clusters that were identified by this cluster analysis were then analysed in excel using a chi square analysis. This procedure was carried out for all of the OTS data analysis chapters.

6 An analysis of the OTS dataset for In-Depth Accidents

6.1 Introduction

Research to improve road traffic safety and reduce casualties has historically concentrated on identifying single factors that cause accidents and their effects. Road safety management in many countries is increasingly using systems approaches to provide further reductions in traffic casualties.

As (Clarke et al., 2005, p. 721) pointed out "In-depth studies of behavioural factors in road accidents using conventional methods are often inconclusive and costly" thus a good understanding of the nature of the data and analysis required is necessary to be able to deduce the sequence and causal links within the accident process.

New intelligent technologies are rapidly being introduced to the road and vehicle environment with the purpose of improving safety and transport efficiency. According to Ljung Aust (2010) the goal of these preventive safety functions, or advanced driving assistance systems (ADAS), is to prevent accidents from occurring and/or to reduce accident severity, by either alerting the driver to potential hazards or by taking over the driving task to some extent, using, for example, autonomous braking in emergency situations. With the increasing development and implementation of these systems within vehicles it is necessary to thoroughly understand the critical situations that can be addressed with the different current and emerging technologies. Accident causation research allows for this analysis by identifying the key factors, human functional failures and interactions that result in a traffic accident.

Detailed accident studies have been carried out in the UK (Carsten, Tight, Southwell, & Plows, 1989; Sabey & Staughton, 1975), but typical accident scenarios have not been developed. Large accident datasets analysed using cluster analysis methods throughout Europe (de Oña et al., 2013b; Depaire et al., 2008) did not include more than one human failure factor or contributory factors, and so the weighting of the cluster reflected vehicular and environmental factors to a greater extent. It was determined that an in-depth look at data coded with accident causation variables would further elaborate on these findings and indicate possible differences if present, as well as allow for an exploration of UK specific scenarios.

The aim of this study is to identify and compare accident causation chains within the OTS database using cluster analysis to identify functional failure sequences with an emphasis on human errors.

6.2 Method

6.2.1 Design

In this chapter a total of two analysis procedures were carried out. Firstly a descriptive analysis of all accident cases collected in the UK OTS study was carried out. Secondly two separate cluster analysis procedures using all multiple vehicle and single vehicle accidents for suitable cases were carried out. The following sections describe how this data was collected, the procedures put in place for this analysis and the cluster analysis.

6.2.2 Sample

Details of the causation factors relating to individual collisions were acquired using in-depth accident data methods on the spot by a group of accident researchers within an average time span of 20 minutes after an accident had occurred. A detailed explanation of the procedures used can be found in the methodology chapter. The results presented in this study are based on 1,614 accidents.

The age distribution from the dataset can be seen in figure 16. Of the 1,877 individuals whose age was coded in the sample the age of the road users were on average 38.0 years old with a standard deviation of 16.2 (figure 16).

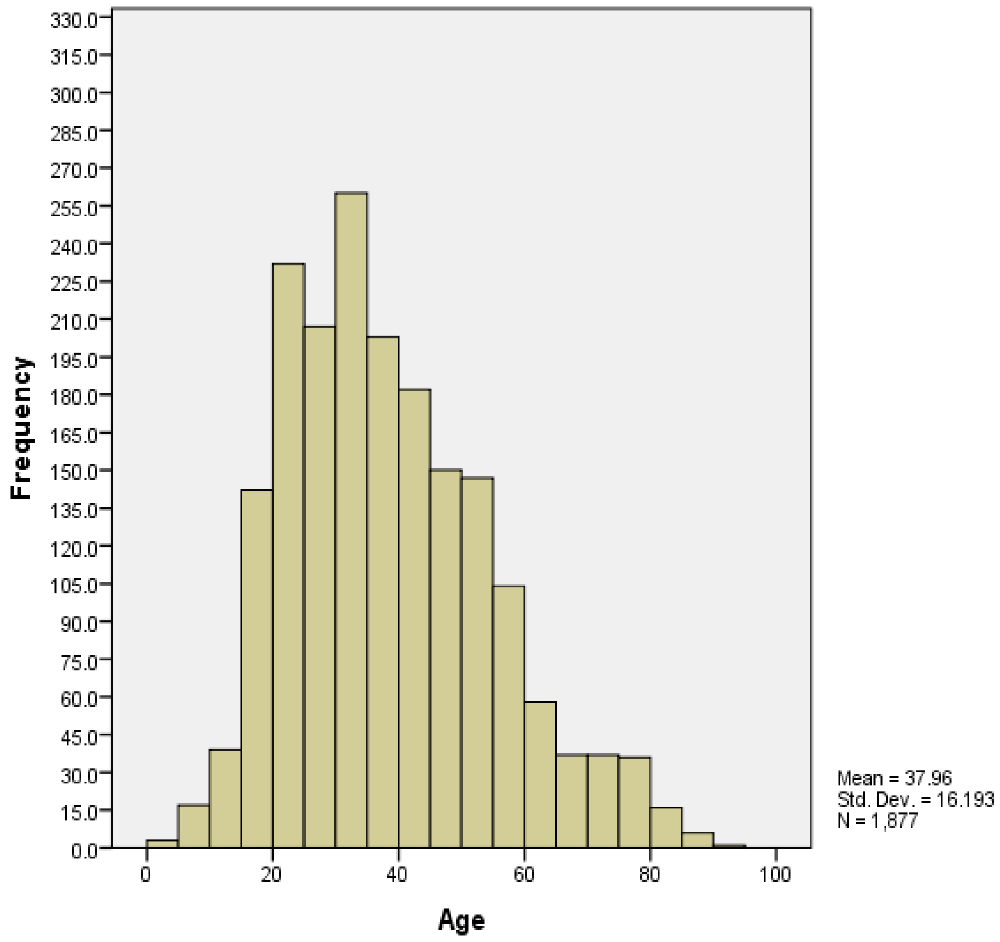


Figure 16: The age distribution for the OTS dataset

6.2.3 Procedure

The attained data contained a total of 1,614 cases, from these cases single vehicle accidents and multiple-vehicle accidents were split into two separate files. There were 543 single vehicle and 1,093 multiple vehicle accidents which included 2,893 interactions. For the cluster analysis all cases that included unknown variables were excluded from the analysis, furthermore four of the single vehicle accidents involved bicycles and were excluded from this analysis on account of the low numbers. For the multiple vehicle accident dataset 22 of the cases did not have the relevant OTS data for the second vehicle and were excluded from this analysis. Two separate cluster analysis procedures were carried out the first for single vehicle accidents (including 366 cases) and the second for multiple vehicle accidents (including 673 cases). Only the first two vehicles that had the main “interaction” within the data were analysed in the multiple vehicle accident cluster analysis study.

6.2.4 Statistical analysis

This study identified similarities between different traffic accidents depending on the specific human error that the road user had made in causing an accident. The sample was analysed by separating the different failure types that the different drivers made during the accident and grouping cases according to these failures.

An identification of key human, vehicular, environmental and infrastructure factors for each of the failure types was then carried out. A descriptive analysis was done to identify all of the manoeuvre types and contributory causation factors within the data. As the dataset used consisted of more than 2,000 variables for each accident it was necessary to use an exploratory analysis tool to find significant factor groupings. Cluster analysis was used in this study as the procedures outputs most closely matched the stated aims and objectives. The basic aim of cluster analysis is to find natural groupings of individuals, In order to carry out cluster analysis the similarity (or dissimilarity) of every pair of individual needs to be measured (Chatfield & Collins, 1980). A table for each cluster highlighting the relevant significant factors were created including the chi square values, variables included, number of cases where the factor was present and the significant values. The number of cases where the factor was present was found by multiplying the percentage and number of cases with the clusters, so sometimes similar values will have different percentages as the values were rounded up. This was the same for all subsequent cluster analysis procedures throughout this thesis.

6.3 Results

6.3.1 Descriptive analysis

An analysis was carried out using all of the cases coded with the HFF method. Relevant demographic, accident type and risk factors were included and broken down with regards to the different failure type that a road user made. A total of 2,893 road user accident data files were analysed and a chi square significance test was carried out for each of the failure types with regards to the coded factor variable coding, which can be seen in table 18. The comparison used the residual value to confirm whether the variable value was over or under-represented according to the overall number. Each value that is over-represented according to this analysis was highlighted in bold in table 18.

In two of the demographic variables (gender and age groups), the accident injury level and the road side speed limit had unknown factors and these values were not included in the chi square analysis. The number of unknown cases was written as a total in the relevant variable row and failure type column in the table. In the variable risk factors the columns do not necessarily have to add up to 100% as each factor can be coded for a road user.

The largest group of accidents coded were prognosis failures (N=1,112). Detection failures (N=626) and diagnosis failures (N=478) were also commonly coded. All of the groups had significant correlations with being the primary road user involved in the accident other than prognosis (expectation of the other road users behaviours) failures, which was the main failure coded for the vehicle that was not contributing to the accidents occurrence. A description for each failure groups significant chi square results is provided below:

Table 18: All road user cases versus failure types with risk factors and other important factors

Factor	Detection N=626	Diagnosis N=478	Prognosis N=1112	Decision N=267	Execution N=158	Overall N=252	N 2893
Gender							
Male	66.8	**74.4**	73.2	**81.7**	68.5	**74.4**	73.2
Female	**33.2**	25.6	**26.8**	18.3	**31.5**	25.6	26.8
Unknown (coding)	56	52	122	43	12	17	302
Vehicle type							

Factor	Detection N=626	Diagnosis N=478	Prognosis N=1112	Decision N=267	Execution N=158	Overall N=252	N 2893
Car	**72.5**	**83.3**	**73.9**	74.2	77.2	**80.2**	75.9
LGV	**6.2**	4.4	**4.8**	2.6	4.4	0.4	4.4
HGV	**8.6**	4.0	**5.4**	2.2	8.2	2.4	5.5
Bus	**1.4**	1.3	**1.3**	0.0	0.0	0.4	1.0
PTW	3.4	5.0	**7.0**	5.6	5.7	4.4	5.5
Cycle	**2.9**	0.6	**2.6**	3.0	1.9	1.6	2.2
Pedestrian	4.5	0.8	2.7	**11.2**	1.3	**10.7**	4.2
Speed limit							
30 mph	**44.9**	33.1	**41.1**	**61.8**	**25.3**	44.4	41.9
40 mph	11.3	12.3	**13.8**	12.0	10.8	12.7	12.6
50 mph	3.4	2.9	**4.3**	3.4	2.5	1.6	3.5
60 mph	15.5	**29.3**	**19.4**	9.4	24.1	19.8	19.6
70 mph	**21.6**	**20.7**	**18.9**	9.0	**36.1**	18.7	19.8
Unknown (coding)	21	8	28	12	2	7	78
Involvement							
Primary	**91.2**	**83.9**	5.1	**97.4**	**91.8**	**100.0**	58.3
Secondary	**5.8**	4.2	**4.0**	3.4	3.8	0.4	4.0
Non contributory	4.5	12.8	**96.5**	4.1	7.0	0.0	40.9
Light Conditions							
Day	**91.1**	70.3	**79.6**	64.0	77.2	55.2	76.8
Night	18.5	**29.7**	**20.4**	**36.0**	22.8	**44.8**	25.2
Injury level							
Fatal	2.4	3.6	**3.5**	5.2	2.5	**8.3**	3.8
Serious	10.9	9.4	**12.5**	**15.0**	13.9	13.5	12.0
Slight	**49.2**	41.0	**50.6**	41.6	27.8	38.1	45.6
Non-Injury	**36.6**	**44.6**	**32.9**	37.8	**55.7**	37.7	37.7
Unknown (coding)	6	7	5	1	0	6	25
Age range							
0-17	3.8	1.7	2.0	**8.2**	3.8	**9.1**	3.6
18-21	5.0	**13.6**	4.2	7.9	4.4	4.0	6.3
22-29	**13.1**	**13.4**	**11.2**	9.0	10.8	10.7	11.7
30-49	**24.8**	21.5	**31.2**	20.6	31.6	21.4	26.4
50-65	**12.1**	7.7	**12.1**	7.1	10.1	8.7	10.5
66+	**5.3**	3.1	3.1	4.1	5.1	**8.7**	4.3
Unknown(coding)	234	190	464	128	58	95	1169
Risk factors							
Alcohol	0.0	0.2	0.2	2.6	1.9	**45.2**	4.4
Speeding	8.0	**42.9**	2.7	**33.0**	**25.3**	**23.8**	16.3
Distraction	**9.7**	1.0	0.1	1.5	**17.1**	4.0	3.7
In a hurry	**25.1**	**25.5**	2.3	**40.8**	10.8	**19.8**	16.6

Factor	Detection N=626	Diagnosis N=478	Prognosis N=1112	Decision N=267	Execution N=158	Overall N=252	N 2893
Breaking the law	**24.4**	**19.2**	0.9	**84.3**	8.2	18.7	18.7
Visibility	**26.2**	3.6	2.5	9.4	3.8	4.8	8.7
Accident type							
Overtaking/ Lane change	**29.7**	27.6	**33.2**	**34.8**	20.9	25.4	30.3
Loss of control	5.1	**17.6**	**12.3**	8.2	**43.7**	**31.3**	14.6
Rear-end	**27.5**	7.9	**24.0**	5.2	8.2	2.8	17.7
Head on	1.4	**6.1**	**7.5**	4.5	5.1	3.6	5.2
Pedestrian	**7.2**	1.3	**5.2**	**9.4**	1.9	**7.9**	5.4
Right turn against	**5.6**	4.8	**6.4**	5.2	0.0	2.8	5.2
Turning	**9.1**	4.6	**8.5**	8.6	0.6	3.2	7.1
Intersection(not turning)	4.3	1.3	**4.9**	**10.9**	2.5	1.6	4.3
Merging	**4.2**	3.8	**3.6**	3.0	0.0	0.4	3.2
Other	5.9	**25.1**	0.0	10.1	**17.1**	**21.0**	9.1

Detection failures: Association between gender and detection failures were significant ($\chi^2=5.23$, df=1, p=<0.05), with females being over-represented in this accident type. All road user types other than PTW riders and pedestrians were more likely to exhibit detection failures ($\chi^2=20.41$, df=6, $p<0.01$). The speed limits ($\chi^2=12.5$, df=3, $p<0.01$) were significant with 30 mph and 70 mph accounting for these values. These accidents most commonly occurred during the day ($\chi^2=5.23$, df=1, $p<0.01$) and had either a slight or non-injury accident as being significant ($\chi^2=9.59$, df=3, $p<0.01$). The risk factors visibility issues, distraction, breaking the law and being in a hurry were highlighted for this failure type ($\chi^2=26.88$, df=5, $p<0.001$). Overtaking, rear-end and pedestrian accidents were also identified as being more likely to occur than the other factors with regards to the total accident data, ($\chi^2=25.89$, df=10, $p<0.01$).

Diagnosis failures: Male road users were more likely to make this failure type and were slightly over-represented ($\chi^2=4.78$, df=1, $p<0.05$). Road users driving cars ($\chi^2=14.84$, df=6, $p<0.01$) as well as roads with high speed limits ($\chi^2=12.31$, df=4, $p<0.01$), were also significant. These failure types occurred more prominently during the night ($\chi^2=4.84$, df=1, p<.0.05), and the main risk factors were speeding, being in a hurry and breaking the law ($\chi^2=15.90$, df=6,

$p<.0.05$). The main accident types for these failures were loss of control and head on accidents, ($\chi^2=25.35$, df=9, $p<0.01$).

Prognosis failures: All road users except for pedestrians were highlighted by the analysis ($\chi^2=24.12$, df=7, $p<0.01$). Due to the large number of cases most of the factors were highlighted as significant. The main factors that were not significant were being the primary contributory vehicle, as well as high risk factors.

Decision failures: Male road users were identified as effecting the significance of the gender variable ($\chi^2=4.07$, df=1, $p<.0.05$). Overtaking/lane changing accidents, pedestrian accidents and intersection accidents were more likely to occur ($\chi^2=24.21$, df=9, $p<0.01$). The main contributory factors were breaking the law, being in a hurry and distraction ($\chi^2=22.24$, df=6, $p<0.01$).

Execution failures: Loss of control accidents as well as the contributory factors distraction and being in a hurry were highlighted ($\chi^2=19.51$, df=9, $p<0.05$) for this failure type.

Overall failures: Being a male road user ($\chi^2=4.37$, df=1, $p<0.05$), and a car driver or pedestrian were more likely for this failure ($\chi^2=14.68$, df=6, $p<0.05$). These failures were more likely to be coded for night time accidents ($\chi^2=5.60$, df=1, $p<.0.05$) compared to the other accident types. Overall failures were more likely to be fatal accidents ($\chi^2=10.33$, df=3, $p<.0.05$), and have younger (ages 0-17) or older (aged 66 or above) road users ($\chi^2=15.70$, df=5, $p<0.01$). Alcohol, speeding and being in a hurry were the most likely risk factors ($\chi^2=113.08$, df=6, $p<0.001$). Loss of control accidents and pedestrian accidents were the most likely accident types ($\chi^2=21.48$, df=9, $p<.0.05$).

Single vehicle accidents

There were a total of 539 single vehicle accidents in the OTS database between the years 2000 to 2003. Table 19 illustrates the number of single vehicle accidents that were collected in the OTS data by vehicle type. The vehicle that had the highest proportion of single vehicle accidents were cars (85.9%) and the second highest were PTWs (7.4%).

Table 19: Vehicle types for single vehicle accidents

Vehicle Type	N	Percent
Car	463	85.9
Light Good Vehicles	12	2.2
Heavy Goods Vehicles/Bus	24	4.5
PTW	40	7.4
Total	539	100.0

The different failure types for all single vehicle accidents in relation to injury severity are illustrated in Table 20. Most of the cases in the dataset are non-injury accidents (58.5%). The most prominent failure types with regards to single vehicle accidents were diagnosis failures (35.6%) and overall failures (26.7%), which together accounted for close to two thirds of all single vehicle accidents. Execution failures (17.1%) were also a prominent group of failures for this accident type. Overall failures (9.4%) had the highest number of fatal accidents and decision failures (13.3%) had the highest number of serious injuries for the analysed cases.

Table 20: Failure types in single vehicle accidents by injury

	Failure type						
Severity	Detection	Diagnosis	Prognosis	Decision	Execution	Overall	Total
Fatal	0.0	2.2	0.0	5.0	1.1	9.4	4.0
Serious	6.8	12.4	0.0	13.3	9.8	9.4	10.6
Slight	29.5	26.3	33.3	13.3	30.4	30.2	26.9
Non-injury	63.6	59.1	66.7	68.3	58.7	51.1	58.5
Total	8.2	35.6	1.1	11.3	17.1	26.7	539

6.3.2 Single vehicle accident cluster analysis

In table 21 all of the factors that were entered into the single vehicle accident cluster analysis are outlined. A total of 13 specific variables were selected according to the most relevant risk factors that are present in single vehicle accidents. They were divided into four groups. The human factors selected were the main failure that the road user was coded as making as well as the main two contributory factors that were coded as occurring in the accident. The contributory factors were grouped according to the contributory human factors sub groups outlined in section 4.2.2. In cases where the failure was descriptive enough and no other factors were necessary 'no factor coded' was selected as the contributory factor. The age group of the road user, the gender type and the vehicle type were also coded. With regards to environmental and infrastructure factors the road type, speed limit, and carriageway class were entered into the analysis. To identify the accident situation the manoeuvre of the road user was also included in the analysis as well as the LAB accident type scenarios that were most prevalent. A detailed list of all of the values counts and percentages can be found in Appendix B (pp. 330).

Table 21: Variables used in the single vehicle accident cluster analysis

Variable	Aspect	Level	Value
Speed limit	Environmental	Accident	≥ 30 mph; 40-50 mph; 60-70 mph
Road area	Environmental	Accident	Urban; Rural
Light conditions	Environmental	Accident	Day; Night
Road type	Environmental	Accident	A class; B class; Motorway; Minor
Failure mechanism	Traffic accident	Road user	Detection; Diagnosis; Prognosis; Decision; Execution; Overall
Gender	Road user	Road user	Male; Female
Age group	Road user	Road user	0-21; 22-29; 30-49; 50-65; 66+
Contributory factor 1	Accident	Road user	Impairment; Alcohol; Psychological factors; Speed; Breaking the law; Inexperience; Distraction; Road Condition; Other road factors; Visibility; Obstacle in road; Vehicular factor; No factor coded
Contributory factor 2	Accident	Road user	Impairment; Psychological factors; Speed; Risk taking; Inexperience; Distraction; Environment; Other factors; No factor coded
Emergency manoeuvre	Accident	Road user	Brake; None; Steered
Road user vehicle type	Vehicle	Road user	Car; PTW, Other
Manoeuvre	Road user	Vehicle	Going ahead; Left bend; Right bend; Intersection; Other
Accident type	Accident	Road user	Leaving lane left; Leaving lane right; Rollover; Collision with obstruction/Hit parked car; Roundabout; Other

Goodness of fit

The latent class cluster analysis focused on the documented 367 single vehicle accident files. The comparison of the BIC and AIC for the different clusters analysed highlighted a 2 (11509.48) and 6 (10747.41) cluster solution respectively, the AIC criterion identified the 6 cluster solution as having the best separation and statistical significance for analysis purposes. The levels for both the AIC and BIC for 2 to 15 cluster solutions can be seen in figure 17.

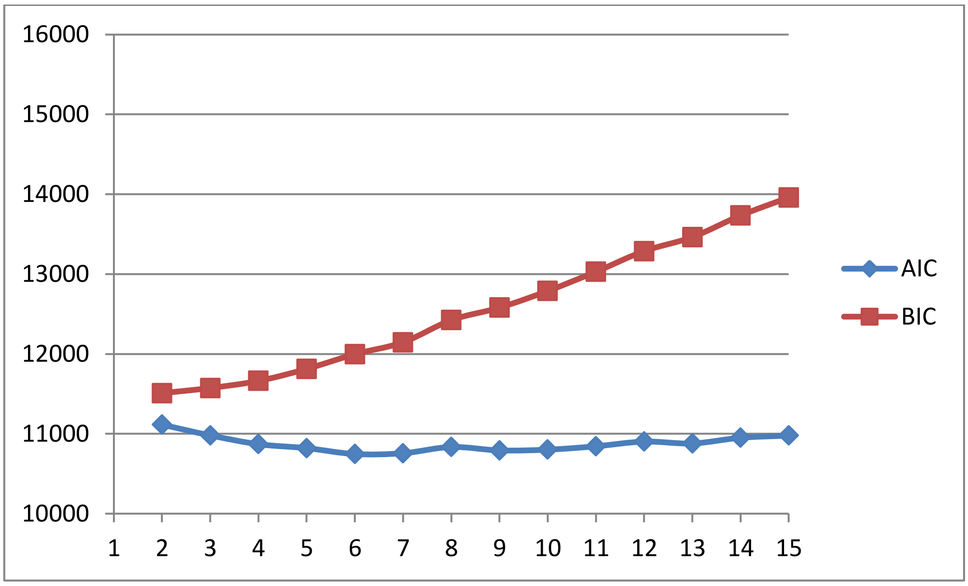


Figure 17: AIC and BIC values for the single vehicle accident cluster analysis

Six distinctive (separated) accident classes were identified resulting in a 6 cluster solution. The results from the full cluster analysis can be seen in Appendix B (pp. 332), factors that were identified as being significantly over-represented are highlighted in bold in that table.

Cluster results

Figure 18 highlights the different number of cases for the single vehicle cluster analysis. The clusters were arranged in size order from cluster 1 through to cluster 6.

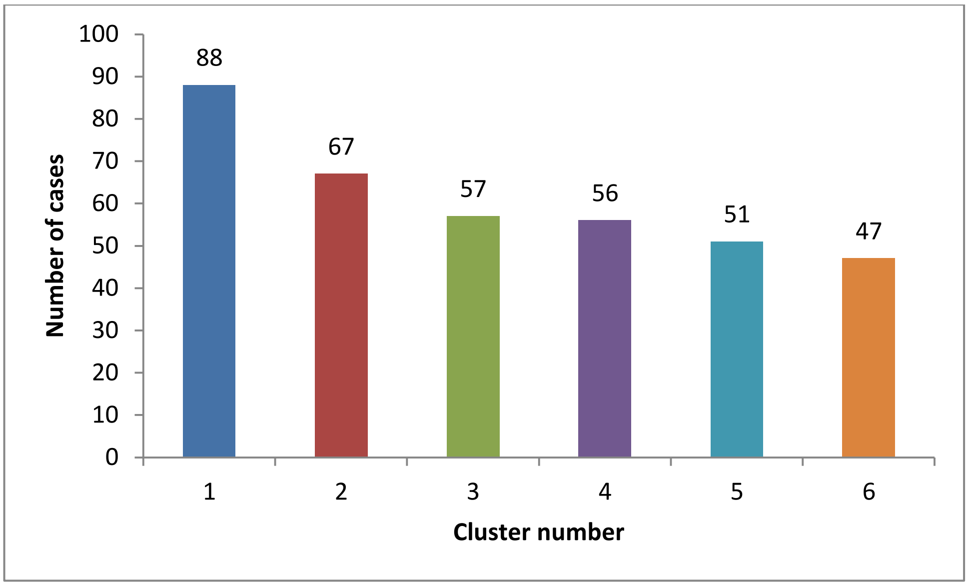


Figure 18: Single vehicle accident cluster sizes

Cluster analysis results

Cluster 1 (n=88)

"Leaving the lane on a bend as a result of speeding"

Table 22 highlights the results for cluster 1, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 22: Single vehicle cluster 1 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Gender	Male	81.5	1	72	4.0	0.05
Age	0-21	29.5	4	26	14.4	0.01
Failure mechanism	Diagnosis	91.5	5	80	130.6	0.001
Contributory factor 1	Speed	88.3	12	78	191.7	0.001
Contributory factor 2	Psychological	36.3	8	32	31.7	0.001
Speed limit	60-70 mph	75.2	2	66	18.8	0.001
Area type	Rural	71.7	1	63	9.4	0.01
Manoeuvre	Left bend	41.7	4	37	107.8	0.001
	Right bend	41.6	4	37		

Accident type	Leaving lane left	53.7	5	47	13.2	0.05

Human Factors

The demographic variables relating to the road user were male road users (81.5%) and the age group between 0 to 21 years old (29.5%). The road user made a diagnosis failure (91.5%), with speed (88.3%) being the first contributory factor and psychological factors (36.3%) the second contributory factor.

Vehicular Factors

No vehicle type was over-represented in this cluster.

Environmental/Infrastructural Factors

The accidents occurred in a rural (71.7%) area on a 60 – 70 mph speed limit (75.2%) road. The road users were on a left hand (41.7%) or right hand (41.6%) bend and left the traffic lane to the left (53.7%) of the road.

Cluster 2 (n=67)

"Leaving the lane due to human or environmental factors"

Table 23 highlights the results for cluster 2, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 23: Single vehicle cluster 2 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Gender	Female	48.9	1	33	20.4	0.001
Age	22-29	51.7	4	35	41.3	0.001
Failure mechanism	Detection	26.4	4	18	61.9	0.001
Contributory factor 1	Psychological factors	26.5	12	18	81.1	0.001
	Road condition	22.1	12	15		
	Visibility	6.6	12	4		

	Obstacle in road	5.8	12	4		
Emergency manoeuvre	Steered	39.0	2	26	23.6	0.001
Vehicle type	Car	92.2	2	62	6.9	0.05
Speed limit	60-70 mph	78.7	2	53	20.0	0.001
Area type	Rural	87.9	1	59	30.8	0.001
Manoeuvre	Going ahead	50.0	4	33	34.7	0.001
Accident type	Leaving lane left	57.5	5	38	12.2	0.05

Human Factors

Female (48.9%) road users were over-represented for this group, and in the age range 22-29 (51.7%). The main failure detection failure (26.4%) was also over-represented for this group. A number of contributory factors were coded for the first contributory factor with the largest group being psychological factors (26.5%).

Vehicular Factors

The vehicle was a car (92.2%).

Environmental/Infrastructural Factors

These accidents occurred in a rural area (87.9%) on a 60 to 70 mph (48.7%) speed limit road. The road users were going ahead (50.0%) and leaving the lane to the left (57.5%) of the road.

Cluster 3 (n=57)

"Accidents occurring due to impairment and alcohol"

Table 24 highlights the results for cluster 3, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 24: Single vehicle cluster 3 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Age	50-65	21.3	4	12	56.3	0.001
	66+	21.9	4	12		

Failure mechanism	Decision	33.4	5	20	92.0	0.001
Contributory factor 1	Impairment	37.9	12	22	202.5	0.001
	Alcohol	38.2	12	22		
Contributory factor 2	No factor coded	54.9	8	3	29.7	0.001
Area type	Urban	54.4	1	31	4.0	0.05
Manoeuvre	Going ahead	39.7	4	23	10.3	0.05
Accident type	Leaving lane right	37.7	5	22	12.5	NS

Human Factors

The significant age group were road users above the age of 50 (43.2% in total). The main failure for this cluster was a decision failure (33.4%). The first contributing factor was impairment (37.9%) or alcohol (38.2%) and the second factor was no factor coded (54.9%).

Vehicular Factors

No vehicle type was over-represented in this cluster.

Environmental/Infrastructural Factors

These accidents occurred on an urban road (54.4) where the road users were going ahead (39.7%). The scenario leaving the lane to the right (37.7%) had a significant residual value despite the accident type group not having a significant chi square value.

Cluster 4 (n=56)

"Accident due to being in a hurry and detection failures in an intersection"

Table 25 highlights the results for cluster 4, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 25: Single vehicle cluster 4 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Age	66+	13.9	4	8	11.0	0.05
Failure mechanism	Detection	24.9	5	14	39.8	0.001
Contributory factor 1	Psychological factors	34.6	12	19	46.8	0.001
	Distraction	9.1	12	5		
	Visibility	9.3	12	5		
Contributory factor 2	Distraction	5.4	8	3	10.3	NS
Emergency manoeuvre	Brake	31.6	2	18	12.6	0.001
Vehicle type	PTW	30.7	2	17	35.0	0.001
Speed limit	30 mph and under	73.2	2	41	65.9	0.001
	40-50 mph	20.6	2	12		
Road type	Minor	46.1	3	26	17.1	0.001
Area type	Urban	90.5	1	51	62.7	0.001
Light conditions	Day	81.3	1	45	12.3	0.001
Manoeuvre	Intersection	48.8	4	37	27.9	0.001
Accident type	Roundabout	28.1	5	16	38.0	0.001

Human Factors

The road users age was 66 years and older (13.9%). The failures that were significant in this cluster were detection failures (24.9%). The first contributory factor was psychological factor (34.6%) or visibility issues (9.3%) and the factor distraction (5.4%) had a significant residual value despite the second contributory factor not having a significant chi square value.

Vehicular Factors

PTWs (30.7%) were over-represented in this cluster.

Environmental/Infrastructural Factors

These accidents occurred in urban areas (90.5%), during the day (81.3%), on a minor road (46.1%), in a 30 mph or under (73.2%) or 40-50 mph (20.6%) speed limit road. These accidents occurred on an intersection (48.8%) and the accident type coded were roundabout conflicts (28.1%).

Cluster 5 (n=51)

“Diagnosis failures in a high speed setting due to vehicle factors”

Table 26 highlights the results for cluster 5, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 26: Single vehicle cluster 5 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Age	30-49	70.5	4	36	32.8	0.001
Failure mechanism	Diagnosis	91.5	5	47	66.8	0.001
Contributory factor 1	Inexperience	10.3	12	5	144.3	0.001
	Road condition	16.5	12	8		
	Vehicle factors	29.0	12	15		
Contributory factor 2	No factor coded	44.3	8	23	11.9	NS
Emergency manoeuvre	Brake	34.1	2	17	14.9	0.001
Vehicle type	Other	22.9	2	12	22.6	0.05
Speed limit	60-70 mph	89.5	2	46	28.5	0.05
Road type	Motorway	44.3	3	23	48.7	0.05
Area type	Rural	93.4	1	48	31.0	0.05
Manoeuvre	Going ahead	51.3	4	26	36.2	0.001
	Right bend	30.5	4	16		

Human Factors

Gender was not significantly over-represented in this cluster and the road user age range 30-49 year olds (70.5%) was significant. The main failure was a diagnosis failure (91.5%). The first contributory factors were vehicle factors (29.0), inexperience (10.3%), or the road condition (16.5%). A braking manoeuvre was made in 34.1% of these cases.

Vehicular Factors

The road user’s vehicle type coded as other (22.9%), which includes LGV, HGV and buses, were over-represented in this cluster.

Environmental/Infrastructural Factors

These accidents occurred in a rural area (93.4%) with a speed limit of 60-70 mph (89.5%) on a motorway (44.3%). These accidents occurred either with the road user going ahead (51.3%) or on a right hand bend (30.5%).

Cluster 6 (n=47)

"Accidents occurring in a low speed setting due to alcohol or breaking the law"

Table 27 highlights the results for cluster 6, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 27: Single vehicle cluster 6 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Road user gender	Male	87.4	1	41	5.6	0.05
Road user age group	0-21	47.9	4	23	33.6	0.001
Road user failure mechanism	Decision	33.4	5	16	73.5	0.001
Road user contributory factor 1	Alcohol	40.0	12	19	129.8	0.001
	Breaking the law	23.3	12	11		
	Inexperience	13.8	12	6		
Road user contributory factor 2	Speed	48.2	8	23	58.5	0.001
Vehicle type	Car	97.6	2	46	8.0	0.05
Speed limit	30 mph and under	83.9	2	39	64.3	0.001
Road type	Minor	50.4	3	24	15.7	0.01
Area type	Urban	82.2	1	39	35.0	0.001
Light conditions	Night	82.3	1	39	40.4	0.001
Accident type	Collision with obstruction/Hit parked car	14.0	5	7	34.7	0.001
	Roundabout	23.7	5	11		

Human Factors

The gender of the road user was male (87.4%). The road users age range 0 - 21 (47.9%) was over-represented. The main failure coded in this cluster were decision failures (33.4%), the significant first contributing factor was alcohol (40.0%) or breaking the law (23.3%) and the second contributing factor was speeding (48.2%).

Vehicular Factors

The vehicle type was coded as a car (97.6%).

Environmental/Infrastructural Factors

These accidents occurred in an urban area (82.2%) on a 30 mph or under speed limit (83.9%) road during the night (82.3%). The accident type was either a collision with obstruction (14.0%) or roundabout accident (23.7%).

Cluster analysis results by injury outcome

In terms of injury outcomes the cluster with the highest percentage of fatal injuries is cluster 3 and the cluster with the highest percentage of serious injuries is cluster 6. Clusters 2, 4 and 5 had a high number of non-injury cases, and cluster 1 had some serious injury cases but not a large number of fatal injuries. These figures can be seen in table 28.

Table 28: Single vehicle accident injury outcome by cluster

Injury severity	Clusters						
	1	2	3	4	5	6	n
Fatal	3	1	10	0	0	4	18
Serious	16	2	7	9	6	9	49
Slight	30	22	21	24	12	12	121
Non-injury	37	42	19	23	33	21	175
Unknown	2	0	0	0	0	1	3
Total	88	67	57	56	51	47	366

6.3.3 Multiple vehicle accident cluster analysis

Table 29 illustrates the different factors that were entered into the multiple vehicle accidents cluster analysis. A total of 17 specific variables were selected to be entered into this analysis according to the most relevant risk factors that are present in multiple vehicle accidents. These variables were divided into four groups. The human factors selected were the main failure that both of the road users were coded as making, as well as the main two contributory factors that were coded as occurring in the accident for the first road user and the main contributing factor that were coded as occurring for the second road user. These factors were coded using the HFF main categories for road user 1 with regards to both the main failure and contributory factor 1. The main failure for road user 2 included the six failure groups and a separate group called only present, which was added as a number of vehicles were coded as such.

Whether the road user was contributing to the accident, was a secondary contributing road user or not contributing to the accidents occurrence was also coded as the level of involvement. The age group and gender of each road user was also included in the analysis.

The vehicle type was also coded for both vehicles. In terms of the environment and infrastructure different factors that described the road type, speed limit, and carriageway class were entered into the analysis. The manoeuvre of both of the road users was also included in the analysis as well as the accident type scenarios that were most prevalent. A detailed table of all of the values counts and percentages can be found in Appendix B (pp. 335), in this table each overly represented significant factor is presented in bold.

Table 29: Variables used in the multiple vehicle accident cluster analysis

Variable	Aspect	Level	Value
Speed limit	Environmental	Accident	≥ 30 mph; 40-50 mph; 60-70 mph
Road Area	Environmental	Accident	Urban; Rural
Light conditions	Environmental	Accident	Day; Night
Road user 1 failure mechanism	Traffic accident	Road user	Detection; Diagnosis; Prognosis; Decision; Execution; Overall

Road user 2 failure mechanism	Traffic accident	Road user	Detection; Diagnosis; Prognosis; Decision; Execution; Overall; Only present
Road user 1 gender	Road user	Road user	Male; Female
Road user 2 gender	Road user	Road user	Male; Female
Road User 1 age group	Road user	Road user	0-17; 18-21 22-29; 30-49; 50-65; 66+
Road User 2 age group	Road user	Road user	0-17; 18-21; 22-29; 30-49; 50-65; 66+
Road user 1 contributory factor	Accident	Road user	Impairment; Alcohol; Psychological factors; Risk; Speed; Breaking the law; Inexperience; Distraction; Road Condition; Other road factors; Visibility Impaired; Obstacle in road; Vehicular factor; No factor coded
Road user 2 contributory factor	Accident	Road user	Psychological; Identification; Risk taking; Traffic control; Atypical manoeuvres other road user; Illegal manoeuvres other road user; Other factors; Visibility; No factor coded
Road user 1 mode of transport	Road user	Vehicle	Car; LGV; HGV/BUS; PTW; Pedestrian/Cycle
Road user 2 mode of transport	Road user	Vehicle	Car; LGV; HGV/BUS; PTW; Pedestrian/Cycle
Road type	Environmental	Accident	A class; B class; Motorway; Minor
Road user 1 Manoeuvre	Accident	Road user	Going ahead; Intersection; Turning left;; Right turn; Left turn; Overtaking; Other
Road user 2 Manoeuvre	Accident	Road user	Going ahead; Intersection; Turning; Overtaking; Slowing in traffic; Other
Accident type	Accident	Road user	Rear-end; Right turn against; Right turn same direction; Left turn; Merging road; Roundabout; Leaving lane; Pedestrian; Going into the opposite lane; Overtaking; Other

The latent class cluster analysis focused on the documented 673 accident files. The comparison of the AIC and BIC for the different clusters highlighted an 8 (25468.12) and 2 (27713.48) cluster solution respectively. The AIC criterion identified the 8 cluster solution as having the best separation and statistical significance for analysis purposes. The levels for both the AIC and BIC for 2 to 15 cluster solutions can be seen in figure 19.

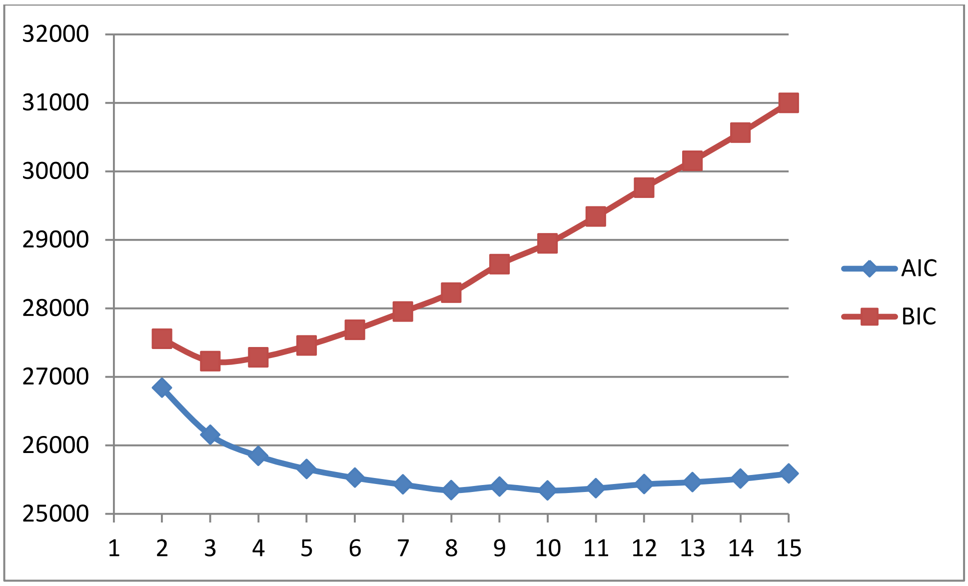


Figure 19: AIC and BIC values for the multiple vehicle accident cluster analysis

Eight distinctive (separated) accident classes were highlighted resulting in an 8 solution cluster. The clusters were ordered with regards to case sizes and the sizes of each of the clusters can be seen in figure 20. A detailed list of all of the cluster results can be found in Appendix B (pp. 338). Factors that were identified as being significantly over-represented are highlighted in bold in the table.

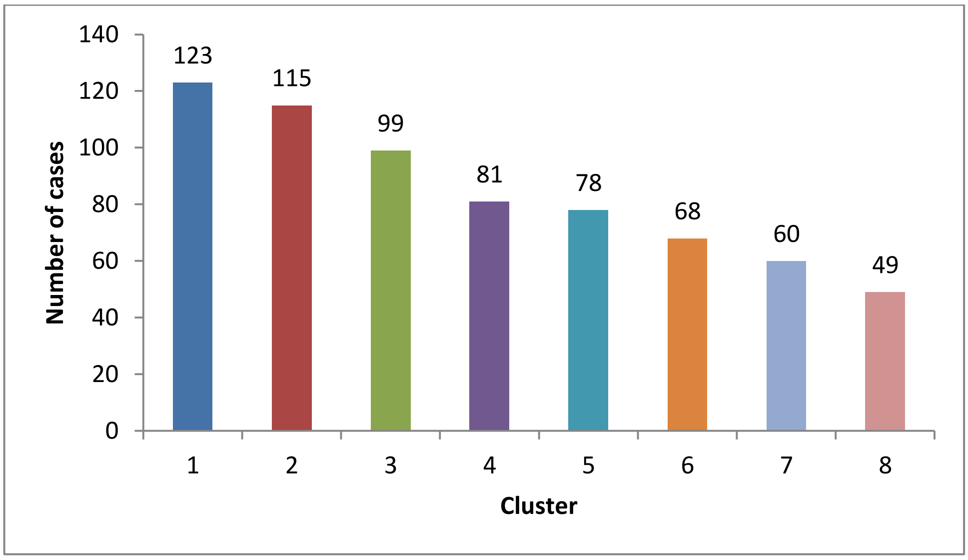


Figure 20: Multiple vehicle accident cluster sizes

Cluster analysis results

Cluster 1 (n=123)

“Turning accidents in a low speed setting due to detection issues from visibility or lane violations”

Table 30 highlights the results for cluster 1, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 30: Multiple vehicle accident cluster 1 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Road user 1 gender	Female	38.7	1	48	5.8	0.05
Road user 1 age group	30-49	49.1	5	60	20.5	0.01
Road user 1 failure mechanism	Detection	65.2	5	80	34.1	0.001
Road user 1 contributory factor	Breaking the law	49.9	13	61	133.2	0.001
	Visibility	20.4	13	25		
Road user 2 age range	0-17	8.9	5	11	22.2	0.001
	18-21	11.9	5	15		
Road user 2 failure mechanism	Prognosis	93.6	6	115	15.0	0.05
Road user 2 contributory factor	Risk taking	5.6	9	7	98.1	0.001
	Illegal manoeuvres other driver	57.0	9	70		
Road user 1 mode of transportation	Car	90.9	4	112	34.9	0.001
Road user 2 mode of transportation	PTW	35.0	4	43	113.9	0.001
Speed limit	30 mph and under	74.0	2	91	32.0	0.001
Area type	Urban	75.0	1	92	7.0	0.01
Road type	B class	23.8	3	29	35.4	0.001
	Minor	41.8	3	51		
Road user 1 manoeuvre	Intersection	77.6	5	95	164.2	0.001
Road user 2 manoeuvre	Going ahead	82.2	5	101	58.7	0.001
Accident type	Right turn against	37.5	10	46	4292.2	0.001

Right turn same direction	19.7	10	24
Left turn	8.1	10	10

Human Factors

Road user 1: Female road users (38.7%) and the age range 30-49 (49.1%) were over-represented in this cluster. The failures for vehicle 1 were detection failures (65.2%). The first contributing factor for road user 1 was breaking the law (49.9%) or visibility issues (20.4%)

Road user 2: The age groups 0-17 (8.9%) and 18-21 (11.9%) were over-represented in this cluster. The failure type that was significant for this road user was prognosis failures (93.6%). The contributing risk factor was coded as other road user's illegal manoeuvres (57.0%).

Mode of transportation

The vehicle coded for road user 1 was a car (90.9%). PTWs (35.0%) were significantly over-represented for the second road user.

Environmental/Infrastructural Factors

The accidents occurred in an urban (75.0%) area. The road type was either a B class (23.8%) or minor (41.8%) road that had a 30 mph or under speed limit (74.0%). The manoeuvre coded for road user 1 was at an intersection (77.6%) and the second road user was coded as going straight ahead (82.2%). The accident type was coded as right turn against (37.5%) or right turn same direction (19.7%).

Cluster 2 (n=115)

"Rear-end accidents in a high speed setting due to detection issues"

Table 31 highlights the results for cluster 2, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 31: Multiple vehicle accident cluster 2 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Road user 1 age group	22-29	26.1	5	30	23.3	0.001
Road user 1 failure mechanism	Detection	86.8	5	100	104.3	0.001
Road user 1 contributory factor	Distraction	15.3	13	17	108.8	0.001
	Potential risk	10.1	13	12		
	No factor coded	28.9	13	33		
Road user 2 failure mechanism	Prognosis	90.5	6	104	7.6	NS
Road user 2 contributory factor	Other road factors	8.5	9	4	122.6	0.001
	Obstacle in road	8.6	9	2		
Road user 1 mode of transport	LGV	11.4	4	13	44.9	0.001
	HGV/BUS	12.9	4	15		
Road user 2 mode of transport	Car	80.4	4	84	10.0	0.05
Speed limit	60-70 mph	55.9	2	97	44.6	0.001
Urban rural	Rural	45.4	1	84	6.1	0.001
Light conditions	Day	87.9	1	13	8.1	0.01
Carriageway	A class	65.5	3	15	40.1	0.001
	Motorway	15.4	3	84		
Road user 1 manoeuvre	Going ahead	61.1	5	13	43.7	0.001
Road user 2 manoeuvre	Intersection	28.6	5	64	233.3	0.001
	Slowing in traffic	44.0	5	52		
Accident type	Rear-end	87.0	10	101	439.3	0.001

Human Factors

Road user 1: Neither gender was over-represented for this cluster. The age group 22-29 (26.1%) was significant. The main failure that this road user made was detection failures (86.8%). The contributing factor for road user 1 was no factor coded (28.9%) or distraction (15.3%).

Road user 2: Gender and age range values were not significant for road user 2. The main road user failure was prognosis failures (90.5%) and the contributing risk factor was coded as other road factors (8.5%) or obstacle in road (8.6%).

Mode of Transportation

Road user 1 was over-represented as an LGV (11.4%) or HGV/BUS (12.9%) driver while road user 2 was a car driver (80.4%).

Environmental/Infrastructural Factors

These accidents occurred in rural areas (45.4%) in a 60-70 mph (55.9%) speed limit road during the day (87.9%) in an A class road (65.5%) or motorway (15.4%). The manoeuvre for road user 1 was going ahead (61.1%) and road user 2 was either at an intersection (28.6%) or slowing down in traffic (44.0%). The accident type was a rear-end accident (87.0%).

Cluster 3 (n=99)

"Urban road low speed accidents"

Table 32 highlights the results for cluster 3, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 32: Multiple vehicle accident cluster 3 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Road user 1 failure mechanism	Prognosis	16.3	5	16	37.1	0.001
Road user 1 contributory factor	Psychological factors	37.7	13	37	59.8	0.001
	Other road factors	6.1	13	6		
Road user 2 failure mechanism	Diagnosis	16.6	6	16	52.5	0.001
	Decision	8.0	6	8		
	Only Present	3.7	6	4		
Road user 2 contributory factor	Psychological	6.0	9	6	44.5	0.001
	Risk taking	6.3	9	6		
	Traffic control	6.5	9	6		

Road user 1 mode of transportation	PTW	15.3	4	15	51.6	0.001
Speed limit	30 mph and under	71.8	2	71	29.4	0.001
Area type	Urban	82.2	1	81	15.7	0.001
Light conditions	Day	86.1	1	85	24.6	0.05
Road type	B class	23.5	3	23	25.1	0.001
	Minor	40.9	3	40		
Road user 1 manoeuvre	Turning right	26.2	5	26	102.5	0.001
	Turning left	11.5	5	11		
	Intersection	15.5	5	15		
Road user 2 manoeuvre	Turning	23.3	5	23	89.5	0.001
	Overtaking	17.7	5	18		
Accident type	Merging road	14.2	10	14	119.4	0.001
	Overtaking	17.7	10	18		
	Other	17.9	10	18		

Human Factors

Road user 1: Gender and age range values were not significant for this cluster. The failures that were over-represented for road user 1 were prognosis failures (16.3%). The first contributing factor for road user 1 were psychological factors (37.7%).

Vehicle 2: Gender and age range values were not significant for this cluster. The failure types that were significant for this road user were diagnosis (16.6%) failures. The contributing risk factor was coded as psychological factors (6.0%), risk taking (6.3%) or traffic control (6.5%) factors.

Mode of transportation

The vehicle coded for road user 1 was a PTW (15.3%). The second road user did not have a significantly over-represented mode of transportation.

Environmental/Infrastructural Factors

These accidents occurred in urban areas (82.2%) in a 30 mph or under road (71.8%) during the day time (86.1%) and in a B class (23.5%) or minor road (40.9%). Road user 1’s manoeuvre was either turning right (26.2%), turning left (11.5%) or intersection (15.5%). Road user 2's manoeuvre was turning (23.3%), or overtaking (17.7%). The accident type was overtaking (17.7%), other (17.9%) or merging road (14.2%).

Cluster 4 (n=81)

"Lane violation due to speed or impairment"

Table 33 highlights the results for cluster 4, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 33: Multiple vehicle accident cluster 4 analysis results

Variable	Value	Percentage	df	N	X^2	Sig.
Road user 1 age group	18-21	20.7	1	17	11.5	0.05
Road user 1 failure mechanism	Diagnosis	42.6	5	34	84.9	0.001
	Prognosis	11.9	5	10		
	Execution	9.5	5	8		
	Overall	15.4	5	12		
Road user 1 contributory factor	Alcohol	8.6	13	7	161.8	0.001
	Speed	42.7	13	35		
	Experience	7.3	13	6		
	Road condition	7.3	13	6		
Road user 2 age group	50-65	29.4	5	24	11.1	0.05
Road user 2 failure mechanism	Prognosis	97.2	6	10	18.8	0.01
Road user 2 contributory factor	No factor coded	65.2	9	53	25.1	0.01
Road user 1 mode of transport	Car	86.7	4	70	26.4	0.001
	Motorcycle	9.2	4	7		
Road user 2 mode of transport	Car	84.4	4	70	19.9	0.001
Speed limit	60-70 mph	46.8	2	38	13.3	0.01
Road area	Rural	58.2	1	47	21.0	0.001
Road type	B class	26.3	3	21	14.2	0.01
Road user 1 manoeuvre	Going ahead	82.6	5	67	94.7	0.001

Road user 2 manoeuvre	Going ahead	83.6	5	68	35.6	0.001
Accident type	Going into opposite lane	67.8	9	55	377.1	0.001

Human Factors

Road user 1: The gender of the first road user was not significant and the age range 18-21 (20.7%) was over-represented. The failures that are over-represented for vehicle 1 were diagnosis failures (42.6%), prognosis failures (11.9%) and overall failures (15.4%). The first contributing factors for road user 1 that predominantly occurred were speed factors (42.7%) or alcohol (8.6%).

Road user 2: Neither gender was significant for the second road user. The age range between 50 to 65 year old (29.4%) was over-represented. The failure type that was significant for this road user was prognosis failures (97.2%). The contributing risk factor was coded as no factors coded (65.2%).

Mode of transportation

The vehicle coded for road user 1 was a car (86.7%). The second road user was also coded as a car driver (84.4%).

Environmental/Infrastructural Factors

The accidents occurred in rural areas (58.2%) with a 60-70 mph speed limit (46.6%). The manoeuvre coded for road user 1 was going ahead (82.6%) and for the second road user was also going ahead (83.6%).The accident type was going into the opposite lane (67.8%).

Cluster 5 (n=78)

“Intersection accidents due to breaking the law”

Table 34 highlights the results for cluster 5, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 34: Multiple vehicle accident cluster 5 analysis results

Variable	Value	Percentage	df	N	X^2	Sig.
Road user 1 age group	50-65	18.6	5	14	18.2	0.01
	66+	13.1	5	10		
Road user 1 failure mechanism	Diagnosis	29.5	5	23	49.1	0.001
	Decision	32.2	5	25		
	Overall	17.4	5	14		
Road user 1 contributory factor	Alcohol	8.8	13	7	123.0	0.001
	Potential risk	12.3	13	10		
	Breaking the law	58.6	13	46		
	Experience	7.6	13	6		
Road user 2 failure mechanism	Prognosis	95.8	6	75	11.1	NS
Road user 2 contributory factor	Illegal manoeuvres other driver	71.4	9	56	110.4	0.001
Road user 1 mode of transport	Car	85.8	4	72	11.0	0.05
Road user 2 mode of transport	Car	82.5	4	67	18.1	0.01
Speed limit	40-50 mph	41.6	2	32	33.6	0.001
Road area	Urban	76.2	1	59	45.1	0.05
Light conditions	Night	31.4	1	25	4.4	0.05
Road type	A class	80.0	3	62	38.4	0.001
Road user 1 manoeuvre	Intersection	56.3	5	44	129.3	0.001
	Turning right	36.5	5	28		
Road user 2 manoeuvre	Intersection	61.0	5	48	90.3	0.001
Accident type	Merging road	37.1	9	29	215.9	0.001
	Roundabout	17.6	9	14		

Human Factors

Road user 1: Road user 1's gender did not have an over-represented group. The age groups 50-65 (18.6%) and 66+ (13.1%) were over-represented for this cluster. The failures that are over-represented for road user 1 were decision failures (32.2%), diagnosis failures (29.5%) or overall failures

(17.4%). The first contributing factor for road user 1 that predominantly occurred was breaking the law (58.6%).

Road user 2: Neither gender nor age groups were significant for this cluster. The failure type that was significant for this road user were prognosis failures (95.8%). The contributing risk factor was coded as other road user's illegal manoeuvres (71.4%).

Mode of transportation

The vehicle coded for road user 1 was a car (85.8%). The vehicle coded for the second road user was also a car (82.5%).

Environmental/Infrastructural Factors

The accidents occurred in an urban area (76.2%). The accident occurred on a road with a 40-50 mph speed limit (41.6%), during the night (31.4%) and in an A class road (80%). The manoeuvre coded for road user 1 was intersection (56.3%) or turning right (36.5%). The second road users manoeuvre was coded as intersection (61.0%). The scenario was a merging road (37.1%) or roundabout (17.6%).

Cluster 6 (n=68)

"Right of way violation due to road user risk taking or illegal behaviour"

Table 35 highlights the results for cluster 6, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 35: Multiple vehicle accident cluster 6 analysis results

Variable	Value	Percentage	df	N	X^2	Sig.
Road user 1 gender	Male	83.6	1	57	26.4	0.001
Road user 1 age groups	30-49	56.1	5	38	49.8	0.001
	66+	11.9	5	8		
Road user 1 failure mechanism	Execution	11.9	5	8	41.4	0.001

Road user 1 contributory factor	Potential risk	12.9	5	9	102.2	0.001
	Vehicle factors	16.4	13	11		
Road user 2 failure mechanism	Diagnosis	14.3	9	10	32.2	0.001
Road user 2 contributory factor	Traffic control	10.0	9	7	59.7	0.001
	Atypical manoeuvres other driver	24.2	9	16		
Road user 1 mode of transport	HGV/BUS	22.4	4	15	62.3	0.001
Road user 2 mode of transport	HGV/BUS	20.0	4	14	43.5	0.001
Speed limit	60-70 mph	89.9	2	61	148.8	0.001
Road area	Rural	88.9	1	60	116.4	0.001
Road type	Motorway	58.1	3	40	271.6	0.001
Road user 1 manoeuvre	Going ahead	51.7	5	35	93.1	0.001
	Intersection	24.6	5	17		
	Other	21.4	5	15		
Road user 2 manoeuvre	Going ahead	79.7	5	54	57.1	0.001
	Overtaking	8.8	5	6		
Accident type	Leaving lane	8.8	9	6	173.1	0.001
	Overtaking	41.3	9	28		

Human Factors

Vehicle 1: The first road users gender was male (83.6%) and age range was 30-49 (56.1%). The main failures that are over-represented for vehicle 1 were execution (11.9%) failures. The first contributing factors for road user 1 that predominantly occurred were potential risk (12.9%), or vehicular factors (16.4%).

Vehicle 2: No gender was significant, while the age group was significant as a chi square analysis but did not have a group that was over-represented. The failure type that was significant for this road user were diagnosis failures (14.3%), and the first contributing factors for road user 2 that predominantly occurred were other road user’s atypical manoeuvres (24.2%)

Mode of transportation

The vehicle coded for road user 1 was a HGV/BUS (22.4%) and for road user 2 was also a HGV/BUS (20.0%).

Environmental/Infrastructural Factors

The accidents predominantly occurred in roads that had a 60-70 mph speed limit (89.9%) in a rural area (88.9%) on a motorway (58.1%). The manoeuvre coded for road user 1 was going ahead (51.7%) and for road user 2 it was also going ahead (79.7%). The scenario was an overtaking (41.3%) or leaving lane (8.8%) accident.

Cluster 7 (n=60)

"Pedestrian accidents occurring as a result of impairment"

Table 36 highlights the results for cluster 7, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 36: Multiple vehicle accident cluster 7 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Road user 1 age group	0-17	68.7	5	41	283.0	0.001
	66+	8.0	5	5		
Road user 1 failure mechanism	Overall	33.5	5	20	54.0	0.001
Road user 1 contributory factor	Impairment	31.8	13	19	126.5	0.001
Road user 2 gender	Female	40.1	13	24	4.5	0.05
Road user 2 failure mechanism	Detection	20.0	6	12	27.8	0.001
Road user 2 contributory factor	Atypical manoeuvres other driver	20.0	9	12	86.0	0.001
	Other factors	8.4	9	5		
	Visibility	25.1	9	15		
Road user 1 mode of transport	Pedestrian/Cycle	100.0	4	60	405.7	0.001
Road user 2 mode of transport	Car	90.3	4	54	14.5	0.01
Speed limit	30 mph and under	81.3	2	49	26.7	0.001

Area type	Urban	85.0	1	51	11.9	0.001
Road type	Minor	57.0	3	34	31.5	0.001
Road user 1 manoeuvre	Going ahead	48.2	5	29	76.5	0.001
	Other	43.4	5	26		
Road user 2 manoeuvre	Going ahead	79.8	5	29	27.2	0.001
Accident type	Pedestrian	83.2	10	50	263.5	0.001

Human Factors

Road user 1: Gender was not significant for this cluster. The age range 0-17 (68.7%) was over-represented. The failures that are over-represented for this road user were overall failures (33.5%) and the contributory factor impairment (31.8%) was also over-represented.

Road user 2: Being female (40.1%) was significant for this road user. The failure type detection failures (20.0%) was significant, and the contributing risk factor was coded as atypical manoeuvres other road user (20.0%).

Mode of transportation

The first road user was a pedestrian/cycle (100.0%). The vehicle type for the second road user was a car (90.3%).

Environmental/Infrastructural Factors

The main factors that were outlined in this analysis were that the accidents occurred on a 30 mph speed limit (81.3%) urban area (85.0%) minor road (57.0%). The accidents that were significant in this cluster were pedestrian accidents (83.2%). Both road users were going ahead (48.2% and 79.8%) and the accident type was a pedestrian (83.2%) accident.

Cluster 8 (n=49)

“Pedestrian/Cyclist to car accidents where road user made the primary contributory behaviour”

Table 37 highlights the results for cluster 8, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 37: Multiple vehicle accident cluster 8 analysis results

Variable	Value	Percentage	df	N	X^2	Sig.
Road user 1 gender	Female	43.4	1	21	4.7	0.05
Road user 1 age group	0-17	14.2	5	7	16.4	0.01
	18-21	20.5	5	10		
	66+	12.7	5	6		
Road user 1 failure mechanism	Decision	44.1	5	22	33.6	0.001
	Overall	14.2	5	7		
Road user 1 contributory factor	Psychological factors	28.7	13	14	16.9	NS
Road user 2 age group	0-17	22.5	5	7	54.1	0.001
	18-21	8.0	5	10		
	66+	8.1	5	6		
Road user 2 failure mechanism	Detection	14.4	6	7	54.2	0.001
	Overall	6.0	6	7		
Road user 2 contributory factor	Psychological	10.2	9	5	37.3	0.001
	Atypical manoeuvres other driver	22.3	9	11		
	Visibility	11.8	9	6		
Road user 1 mode of transport	Pedestrian/Cycle	48.8	4	24	56.8	0.001
Road user 2 mode of transport	Pedestrian/Cycle	48.8	4	24	140.9	0.001
Speed limit	30 mph and under	92.2	2	45	36.8	0.001
Area type	Urban	95.9	1	47	22.6	0.001
Light conditions	Night	34.3	1	17	4.5	0.05
Road user 1 manoeuvre type	Intersection	54.1	5	28	26.5	0.001
Road user 2 manoeuvre type	Other	23.1	5	11	122.7	0.001
	Intersection	52.2	5	27		
Accident type	Pedestrian	91.8	10	44	262.5	0.001

Human Factors

Road User 1: Road user 1's gender was female (43.4%) and the age range 0-17 (14.2%), 18-21 (20.5%) and 66 years or older (12.7%) were significant for this cluster. The failures that are over-represented are decision failures (44.1%) or overall failures (14.2%). The first contributing factor for road user 1 was psychological factors (28.7%)

Vehicle 2: The road users age range was either 0-21 (30.5%) or 66 and above (8.1%). The failure types were detection failures (14.4%) or overall failures (6.0%). The contributing risk factor was coded as other road user's atypical manoeuvres (22.3%) or psychological factors (10.2%).

Vehicular Factors

The vehicle coded for road user 1 was a pedestrian/cycle (48.8%). The second road user was coded as a pedestrian/cycle (48.8%).

Environmental/Infrastructural Factors

The accidents occurred in an urban area (95.9%). The road had a 30 mph speed limit (92.2%). These accidents occurred disproportionally during night time (34.3%). The manoeuvre coded for road user 1 was intersection (54.1%), and the manoeuvres for road user 2 were coded as other (23.1%) or intersection (52.2%). The accident was coded as pedestrian (91.8%).

Cluster by injury

In terms of injury outcomes the cluster with the highest percentage of fatal and serious injuries is cluster 4 which had the highest number of fatal injuries and the second highest number of serious injuries. These figures can be seen in table 38.

Table 38: Multiple vehicle accident injury outcome by cluster

Injury severity	Clusters								
	1	2	3	4	5	6	7	8	n
Fatal	5	1	1	12	0	2	5	3	29
Serious	18	7	17	17	10	10	14	15	108
Slight	84	71	59	40	45	37	40	29	405
Non-injury	16	36	22	12	23	18	1	2	130
Unknown	0	0	0	0	0	1	0	0	1
Total	123	115	99	81	78	68	60	49	673

6.4 Discussion

A latent class cluster analysis was performed on both single and multi-vehicle accident types in order to understand, and statistically identify, the meaningful accident scenarios with regards to these different types of accidents. One of the main points of the analysis was to demonstrate different accident scenario groupings using significant factors. The results of the two analyses are discussed in the sections below.

6.4.1 Descriptive statistics analysis results

Failure type

The analysis of the descriptive statistics shows that the group of accidents that were most often coded were prognosis accidents. These types of accidents occur when a road user can not anticipate a manoeuvre that another road user is making, and so the road user either does not make a mitigating emergency manoeuvre or does not have time to react to the other road user.

In relation to gender females were more likely to make detection failures, execution failures and prognosis failures, while males were more likely to make decision failures and overall failures. These findings are similar to studies that have used macroscopic accident data (Clarke et al., 2007).

In terms of the road type decision failures were most commonly occurring in 30 mph speed limit roads, while execution failures mostly occurred in higher speed roads and under-represented in lower speed roads. The link between speed and crash rate has been well documented (Aarts & van Schagen, 2006; Elvik et al., 2004), so the execution failures may be due to the possibility of recovering from an error made when the speed is low but having a shorter time gap with higher travelling speeds

Contributory factor

Breaking the law (18.6%), speed (16.3%) and in a hurry (16.6%) were the main contributory factors that were identified in the analysis. Violations are a common cause of accidents occurring, in the data these behaviours were particularly over-represented for diagnosis and decision based failures.

The nature of the failures that road users make while speeding leads to issues with diagnosing the road, illegal manoeuvring or, when combined with a decrease in sensory ability, not being able to react to the situation ahead.

The road user's choice led to them either not anticipating the roadway, and to not be able to diagnose the situation, or their rule breaking leading to a conflict situation.

Road user age

In terms of age younger road users (18-21) were over-represented in the dataset for diagnosis failures. Older road users were over-represented in terms of detection and prognosis failure types. These issues point towards male younger road user's higher propensity towards risk taking (Laapotti & Keskinen, 1998).

6.4.2 Single vehicle accidents cluster analysis results

Results for the single vehicle accidents provided a clear definition of scenarios with the analysis classifying the accidents into 6 different clusters. A detailed listing of all relevant factors present in the single vehicle cluster model as well as some explanatory descriptive variables for the clusters is presented in table 39.

The two largest groups of clusters were due to the road user speeding and an individual leaving the lane due to psychological factors or the road condition in a high speed setting. The nature of single vehicle accidents for men was related to speed and impairment resulting from either a decision to make a dangerous manoeuvre or misdiagnosis of the roadside obstacles. The main type of failure for women was detection failures which can be seen in cluster 2. Differences between male and female single vehicle accidents were also observed, with females being 2.5 times more likely than males to make detections failures and 1.5 times more likely to make execution failures. The common accident failure types for males were diagnosis failure, which they were 1.5 times more likely to make compared to females. The majority of cases were identified as loss of control accidents, though there were differentiations based on different group attributes with regards to these accidents.

Cluster 1

Cluster 1 identified a diagnosis failure due to speed in a high speed setting. In terms of gender and age, males and young and middle age road users trended more for these group of accidents. The road user did not make a reactionary manoeuvre during these accidents. Some additional factors that were identified by the cluster analysis were visibility issues, being in a rural area and leaving the road lane.

Cluster 2

Cluster 2 presented single vehicle accidents where psychological factors were related to detection failures. Over half of the cases present in this cluster were vehicles that were going straight and made an emergency steering reaction. These results are in line with the findings of Lapooti and Keskinen (1998), though the contributory factors for female single vehicle road users were cognitive factors in that study.

Cluster 3

Cluster 3 represented an accident type that is related to impairment and alcohol, where the road user did not make an emergency manoeuvre and was travelling on a straight road during the night.

Cluster 4

Of the 36 documented PTW rider failures for single vehicle accidents, 17 were present in cluster 4 with an equal distribution between 4 different failure mechanisms (detection, diagnosis, execution and overall failures).

Cluster 5

Cluster 5 grouped together a number of accidents on a high speed road due to diagnosis failures. In these accidents the road user was not able to correctly diagnose the road setting due to vehicle issues or the road condition.

Cluster 6

Cluster 6 grouped young male riders that either made a violation or took alcohol. The main failures for this group were decision failures.

Table 39: Single vehicle cluster analysis model and variables

Cluster/ N of cases/% of cases	Descriptive information		Cluster model						
	Accident type	Casualty level	Main failures	Contributory factor 1	Contributory Factor 2	Gender/ Age	Manoeuvre	Road setting	Accident setting
1 88 24.0%	Loss of control on a bend	Fatal (0) Serious (6) Slight (12) Non-injury (33)	Diagnosis	Speed	Psychological	Male/ 0-21	Leaving lane left	Left bend Right bend	60-70 mph Rural area
2 67 18.3%	Loss of control on a straight road	Fatal (1) Serious (2) Slight (22) Non-injury (42)	Detection	Psychological Road condition		Female/ 22-29	Leaving lane left	Going ahead	60-70 mph Rural area
3 57 15.6%	Loss of control on a straight road or a bend	Fatal (4) Serious (9) Slight (12) Non-injury (22)	Decision	Alcohol Impairment	No factor coded	50+	Leaving lane right	Going ahead	60-70 mph Urban area
4 56 15.3%	Loss of control in an intersection	Fatal (0) Serious (9) Slight (24) Non-injury (23)	Detection	Psychological	Distraction	66+	Roundabout	Intersection	Minor road 30 mph Urban area Daytime
5 51 13.9%	Loss of control on a bend	Fatal (3) Serious (16) Slight (30) Non-injury (39)	Diagnosis	Vehicle factors Road condition		30-49	Leaving lane left	Right bend	B road 60-70 mph Rural area Daytime
6 47 12.8%	Loss of control on a straight road	Fatal (10) Serious (7) Slight (21) Non-injury (19)	Decision	Alcohol Breaking the law	Speed	Male/ 0-21	Roundabout Other		Minor road 30 mph Urban area Night-time

6.4.3 Multiple vehicle accidents cluster analysis results

The multiple vehicle accidents latent class cluster analysis provided an 8 cluster result. Due to the large amount of data within the analysis most of the cases were characterised by the manoeuvre and accident type that occurred.

A detailed listing of all relevant factors present in the multiple vehicle cluster model as well as some explanatory descriptive variables for the clusters is present in table 40.

The single vehicle accidents identified different scenarios for male and female road users, as well as PTW riders compared to car drivers or pedestrians. The multiple vehicle cluster analysis differentiations were based less on demographic variables. The multiple vehicle accident analysis identified differences between a number of turning accidents with regards to the human failure that was made. Each cluster result is discussed in the section below.

Cluster 1

The largest cluster grouped turning accidents into one group, particularly right turn accidents which in the UK involve crossing the roadway against oncoming vehicles. These accidents were right of way violations where the first road user did not detect the second road user due to the road user disobeying traffic rules in nearly half of these cases. These accident types were evenly distributed over all road user ages and for the first road user females were over-represented and for the second road user males were not over-represented but accounted for 78.1% of the cases. The described gender effect found was similar to Clarke et al.'s (1998) findings.

The structure of the cluster identified that similar accidents occur to all types of road users. PTWs were also over-represented as the second road user in this accident type. Nearly 50% of these accidents involved VRU's, as the second road user that did not expect the first road user to make their manoeuvre. In nearly half of these cases the road user either concentrated on 'visibility constraints' or 'information processing' concentrating on a portion of the accident.

Cluster 2

The second largest group of accidents were rear-end accidents where the road user was either distracted or not paying attention to the road user ahead in high speed situations following a vehicle that braked. The road user ahead was slowing down in traffic and this situation was not identified by the first road user, most commonly due to this vehicle not expecting the vehicle in front to be static or slowing down. Over 80% of all rear-end accidents within the data were present in this cluster. In a large number of these cases a breaking manoeuvre was made but due to the high speed setting the accident occurred. Visibility conditions were not an issue in this accident type rather ‘hurried information acquisition’ or the expectation that the other road user would not be in front of the driver were of particular importance. This study showed a difference compared to Singh (2003) in that the road users did not trend as younger drivers, though 65% of the striking vehicle did involve male drivers similar to the US sample, the difference in age range may be a result of either the smaller sample size or the differences between UK and US drivers.

Cluster 3

Cluster 3 is an amalgamation of different accidents in a low speed limit road. Close to 25% of the road users made a steering avoidance manoeuvre but were not able to stop the accident from occurring. These accidents occurred as a result of manoeuvring by road user 1 and a number of failures were tied to this accident type, the largest of which was detection failures despite this failure type not being over-represented.

Cluster 4

Cluster 4 identified a lane violation accident in a high speed limit road. The results highlighted speed as an important factor and diagnosis failures or overall failures were identified as the most significant failures. For this cluster most of the high injury setting accidents either occurred as a result of detection failures or decision failures. When road users either did not detect the road user ahead or made a decision to undertake risky behaviour injury outcomes were found to be more severe.

Cluster 5

Cluster 5 identified a situation at an intersection where the first road user was either turning or going ahead. These accidents were based on right of way violations due to an incorrect decision or diagnosis of the accident. These accident types had a low number of serious injury accidents, there were mostly slight injury outcomes.

Cluster 6

Clusters 6 highlighted an accident situation where right of way violations occurred due to a detection failure or diagnosis failure, these failures had high descriptive values despite not being statistically over-represented. Road user 2 was making an overtaking manoeuvre and this was not perceived by the first road user. These failures were due to incorrect overtaking manoeuvres with vehicles going in the same directions.

Cluster 7 and 8

Clusters 7 and 8 highlighted two pedestrian accidents where the pedestrian undertook risky behaviours that did not allow the other road user to identify them or react in time. Cluster 8 identified a pedestrian conflict where in nearly half of the cases the pedestrian made the overall failure and in the other half the road user other than the pedestrian made the main failure. These clusters grouped together all pedestrian accident types, and there is a high likelihood that the difference in accident configurations compared to the other clusters, due to pedestrians being present in the accidents, caused this difference. These accident types will be further analysed in detail in chapter 9 when analysing pedestrian accident specific clusters, as the cluster analysis grouping was made based on the different road user type rather than a specific accident.

Table 40: Multiple vehicle cluster analysis model and variables

Cluster/ N or cases/% of cases	Descriptive information			Cluster model					
	Accident type	Configuration	Casualty level	Road user	Main failures	Contributory factor	Gender/ Age	Manoeuvre	Accident setting
1 123 cases 18.3%	Right turn (same direction/against)	Car to Car (56) Car to PTW (38)	Fatal (5) Serious (18) Slight (84) Non-injury (16)	1	Detection	Breaking the law	Female/ 30-49	Turning at intersection	B road/ Minor road 30 Mph Or under speed limit Urban area Night-time
				2	Prognosis	Other driver illegal manoeuvres	Male/ 0-17 18-21	Going ahead at intersection	
2 115 cases 17.1%	Rear-end	Car to Car (70) LGV to Car (10) HGV to car (10)	Fatal (1) Serious (7) Slight (71) Non-injury (36)	1	Detection	In a hurry	Female/ 22-65	Going ahead	A class road/ Motorway 60-70 Mph Rural area Daytime
				2	Prognosis	Road factor	Male/ 30+	Slowing in traffic Intersection	
3 99 cases 14.7%	Urban road low speed accidents	Car to Car (54) Car to PTW (13)	Fatal (1) Serious (17) Slight (59) Non-injury (22)	1	Detection Decision	Psychological factors	22-29	Turning right Other	B class/ Minor road 30 mph Urban area Daytime
				2	Prognosis	No factors coded	Female/ 22-29	Turning Overtaking	
4 81 12.0%	Going into opposite lane	Car to Car (62)	Fatal (12) Serious (17) Slight (40) Non-injury (12)	1	Diagnosis	Speed	Male/ 18-21	Going into opposite lane	B class 60-70 Mph Rural area Daytime
				2	Prognosis	No factor coded	50-65	Going ahead	

5 78 cases 11.6%	Intersection accidents	Car to car (61)	Fatal (0) Serious (10) Slight (45) Non-injury (23)
6 68 10.1%	Overtaking situations	Car to Car (27) HGV/Bus to Car(23) Car to PTW(9)	Fatal (2) Serious (10) Slight (37) Non-injury (19)
7 60 cases 8.9%	Pedestrian Impairment	Pedestrian to car (54)	Fatal (5) Serious (14) Slight (40) Non-injury (1)
8 49 cases 7.3%	Pedestrian Law breaking	Car to pedestrian (24) Pedestrian to car (21)	Fatal (3) Serious (15) Slight (29) Non-injury (2)

1	Decision Diagnosis Overall	Breaking the law	Male/ 50+	Merging road/ Roundabout	A class road 40-50 Mph Night-time Urban area
2	Prognosis	Other road user illegal manoeuvre	Male/ 18-29	Intersection	
1	Detection Execution	Risk factors Vehicle factors	Male/ 30-49	Overtaking Other	Motorway 60-70 Mph Daytime
2	Diagnosis Prognosis	Other road user atypical manoeuvre	22-29	Going ahead	
1	Overall	Impairment	Male/ 0-17	Pedestrian	Minor road 30 Mph Urban
2	Prognosis	Visibility Other road user atypical manoeuvre	Female/ 30-65	Going ahead	
1	Decision	Breaking the law Psychological factors Impairment	Female/ 0-21 66+	Pedestrian	A class road 30 Mph Urban Night-time
2	Detection Decision	Other road user atypical manoeuvre Visibility	Male/ 0-17	Intersection other	

6.4.4 Countermeasure indications

Single vehicle accidents

The single vehicle accident clusters point to speed as being an important factor, both in low speed limit and high speed limit areas. The possible countermeasures related to speed are plentiful and can be selected based on the traffic environments specific requirements. The analysis identified that different factors contribute to the road user's difficulty in analysing the roadside and undertaking appropriate manoeuvres. Females and males were found to have different accident configurations in similar road situations. The possible countermeasure indications for these accidents are outlined below. The factors that can be considered are:

- Speed
- Alcohol
- Road conditions

The relationship between speed, injury and accident occurrence has been well documented throughout the literature (Baruya & Finch, 1994; Elvik et al., 2004). Countermeasures to reduce speed are not just important for single vehicle accidents but for all accident types, as with lower speeds the time that is possible for a countermeasure increases (Carsten & Tate, 2005). Intelligent speed adaption devices have been estimated to reduce up to 33% of accidents on urban roads in field trials (Lai, Carsten, & Tate, 2012).

Simple laboratory task performance has been showed to be highly influenced when they involve secondary impairment (Fillmore, Stockwell, Chikritzhs, Bostrom, & Kerr, 2007; Holloway, 1995). A study by Clarke et al. (2007) showed that alcohol was present in 20% of a sample of 1185 fatal traffic accidents in the UK between the years 1994-2005. Dunaway, Will, and Sabo (2011) listed a large number of possible alcohol prevention measures ranging from relating to individual measures (deterrence laws) to alcohol control policies which can prove effective in the dropping of these accident types.

Roadside factors identified by this study were road defects/layout which was a contributory factor in 6% of the cases and road contaminants which was a factor in 8% of all single vehicle accidents. The condition of the road is particularly important for drivers and depending on the roadside infrastructure a number of possible countermeasures are available, such as geometric countermeasures, signalization countermeasures and road side markings/signs. These countermeasures should be made based on site specific criteria, and though outside the scope of this study are important nonetheless.

Multiple vehicle accidents

In terms of the multiple vehicle accidents the clusters were based around the crash configurations turning accidents, rear-end accident, lane violations, intersection accidents and right of way violation accidents.

The issues identified in turning accidents for cluster 1 were detection based, over 25% of the cases were due to the road user having hurried or ineffective visual search patterns for other road users with conflicting paths and close to 20% were due to visibility constraint conditions. Most of these accidents occurred at intersections (77.2%), where the omission of detection of a key element will lead to the conflict or accident situation to occur (Clarke et al., 1998). Ideally the safety systems that would help these accident types are inter-vehicle communication systems which provide information to the road user depending on other vehicle/VRU’s behaviour and impending conflict situations.

The lane violation accidents from cluster 4 were due to speed or alcohol and the countermeasures for this cluster could be speed based measures or a lane departure warning system.

Intersection accidents identified in cluster 5 were due to breaking the law in an intersection setting. These accidents were due to decision or diagnosis failures. The two types of failures for cluster 6 are making an incorrect diagnosis of the road condition or taking risk due to reaching their desired destination earlier. Training, education and enforcement countermeasures for these accident types are possible, as the failure was mainly connected to road user behaviours as well as the latent conditions producing these failures.

\With regards to the rear-end accidents identified in cluster 2 the main issues were also found to be detection based, with the road user performing in such a way as to not be able to evade hitting the road user in front. It is difficult to find suitable countermeasures for the vehicle that is being struck from behind due to physical limitations and avoidance manoeuvres difficulties. Possible countermeasures should rather concentrate on the issue of distraction or inattention in rear-end accidents in the vehicle that is hitting the rear of the vehicle in front. This issue can be tackled with the use of active safety measures to alert the road user that the length that they are keeping with the vehicle in front is not appropriate. As identified by Davis and Swenson (2006) the vehicles in front longer reaction time to situations which occur in front of this vehicle can in turn cause a shorter reaction time being allotted to the preceding vehicle and in this case a longer following distance would provide a better safety margin, especially considering that in this dataset most of the rear-end accidents occurred on high speed roads.

6.4.5 OTS sampling compared to national data

A study carried out by Richards, Cookson, and Cuerden (2010) linked the OTS and Great Britain national accident (STATS 19) datasets to each other considering all cases collected between the years 2000-2006. This study compared the linked OTS dataset to the STATS19 dataset in terms of injury severity, accident time of day, accident month, road user age and casualties.

In terms of injury outcomes the severity of the accidents in OTS were found to not be representative of the accidents in STATS19, this test did not include the OTS cases that were coded as non-injury. In terms of accident time of day no difference was found between the OTS and STATS19 linked datasets. The month that the accident occurred in the OTS dataset is significantly different to the regional and national STATS19 datasets.

In terms of vehicle types the differences between OTS and the regional and national STATS19 datasets were found to be significant to the 99% level. The OTS dataset had more goods vehicles and PTWs.

There was found to be no significant difference between the OTS and STATS19 data for all casualties, though significant differences in the distribution of the gender of casualties of all severities were found. In these accidents, there were a slightly greater proportion of male casualties in the OTS linked cases.

The differences between OTS and the regional and national STATS19 datasets in terms of road user category were found to be significant. The OTS cases contained a slightly higher proportion of cars, motorcycles and goods vehicles, and a lower proportion of pedestrians and pedal cycles.

This analysis demonstrates that some caution needs to be taken when using OTS data as it is not representative for all accident variables. The analysis carried out was for OTS phases 1 (2000-2003) and phase 2 (2004-2006), in phase 3 (2007-2010) the sampling plan was changed so that the accident cases collected would be more representative of the Great Britain national accident statistics.

6.5 Summary

Chapter 6 described a detailed analysis of the OTS data collected between the years 2000-2003 using descriptive and cluster analysis methods to identify the different failure and interaction sequences found within the accident data.

The descriptive analysis unveiled differences between failure types with regards to demographic variables, risk factors and other accident related variables. This analysis formed the foundation to select the relevant variables for the cluster analysis.

The results from the single vehicle cluster analysis identified a number of loss of control accidents that had different configurations based on the main failure that the road user made and gender and age variables. The multiple vehicle cluster analysis results were based around the manoeuvre and accident configuration. Both of the cluster analysis results were discussed with regards to countermeasure indications.

7 An analysis of the STATS19 dataset

7.1 Introduction

Chapter 6 provided an analysis of in-depth accident data with accident causation variables using multiple and single vehicles. The statistical methodology used was latent class cluster analysis. The current chapter will provide a comparison of latent class cluster analysis methods when using national accident data and perform a comparison of the results from this analysis with the results from chapter 6.

The aims of this study are to compare the differences between microscopic and macroscopic coded data with regards to understanding an accident when using accident causation data and cluster analysis methodology.

7.2 Analysis of failure sequences

In this study an analysis of all STATS19 data with contributory factors was carried out using data from the year 2005 to compare the results from the OTS multiple vehicle study with. This data was selected as the contributory factors data was collected by all police forces starting from 2005 onwards. The data from 2005 was the closest sample to compare with the OTS data collected between the years 2000 to 2003 and thus was used as a comparison.

In this chapter an analysis of all STATS19 data with contributory factors was carried out for the year 2005. The analysis only included two vehicle accidents, which is similar to the study carried out by Depaire et al. (2008). Cluster analysis methods were used for analysis purposes, a detailed description of the methods and procedure used can be found in the methodology chapter.

7.2.1 Participants

The accident data were collected in the year 2005 by police forces around Great Britain. All accidents occurring and reported to the police in Great Britain are collected and inputted either on scene or when the accident is reported. All accidents are then inputted into the STATS19 database which holds records of all reported accidents reported throughout Great Britain. Due to the large number of cases present in the dataset, and for reasons of comparison, it was determined that a comparison of multiple vehicle accidents involving two vehicles would be carried out. This data is the closest to the multiple vehicle accidents analysis carried out on the OTS data. The results presented in this study are based on 55,474 accidents involving two vehicles selected from this dataset.

7.2.2 Procedure

All data used in this study was from the Stats19 national road accident database. This data was provided by the Department for Transport, Great Britain. Each accident case is collected either on scene or retrospectively from reports by police accident investigators using the STATS19 report form. This form includes information on road user, vehicular, infrastructure and environmental factors providing information on these factors as well as factors related to human causality and injury casualty (Pai, 2011).

Accident cases are only collected for accidents where an injury has occurred. The possible injury levels are fatal, serious, and slight injury. Fatal injury is defined where death occurs within 30 days as a result of the accident. A serious injury is any injury that results in the individual requiring medical treatment and most likely staying in a hospital. A slight injury is any injury that is not severe but requires attention on site (Pai, 2011).

7.2.3 Accident causation measures

The accident causation measures data original form was collected starting in 1949, the year that accident data collection was started in the UK. The aim of this form was to allow investigators to identify the factors that they believe contributed to the accidents occurrence. This system was reviewed and improved every five years since its inception, though the form was removed as a data collection requirement after debate about its subjective nature following a review in 1959, despite this in 1994 half of the police forces collecting data used some version of this form (Broughton, 1997).

A report carried out by Maycock (1995) identifying three different systems that the police groups used, persuaded the DfT to commission the development of a contributory systems measure by the Transport Research Laboratory (TRL). This system was based on the work carried out in the ITS study by Carsten et al. (1989) to use a hierarchical method to code the data viewing, though the four level hierarchy was deemed unnecessarily complex and a two level hierarchy was considered in its place (Broughton, 1997; Gkikas, 2009);

1. Precipitating factors; The immediate failures that lead to an accident
2. Contributory factors; The factors for the failures and manoeuvres

The original version allowed for up to three precipitating factors and contributory factors to be coded in decreasing importance, but a review of police accident cases with the form limited the precipitating factor to 1 and allowed the contributory factors to be coded as definite, probable and possible.

The Transportation Research Group in Southampton University provided a review suggesting a revised form for collecting contributory data, though for ease of use a different layout was adopted (Hickford & Hall, 2004). The outcome of that work was the STATS19 contributory factors form now in use, including seventy-six contributory factors and also an option to report “other factor” by text description (Gkikas, 2009). The factors are grouped in five main categories: (1) road environment contributed (nine factors), (2) vehicle defects (six factors), (3) driver/rider only (forty-seven factors), (4) pedestrian only (ten factors), and four factors for special codes (stolen vehicle, vehicle in course of crime, emergency vehicle on call, vehicle door open/closed negligently). The driver/rider category is further subdivided into five subcategories: injudicious action, error or reaction, impairment or distraction, behaviour or inexperience, and vision affected (by). The reporting officer can select up to six factors from the grid, relevant to the accident. Previously suggested three and four-point scales of confidence are now substituted by a simple two-point scale: the officer indicates for each factor whether s/he considers it “very likely” or just “possible”. The system allows for more than one factor to be related to the same road user and for the same factor to be related to more than one road user, if appropriate. This allows the police officer sufficient flexibility to include the necessary details and in a concise manner.

7.2.4 Data handling

The accident data files were provided in 4 separate excel files separating the accident, casualty, contributing factor, and vehicle data files. The accident file was separated by the author into all accidents that involved 2 road users only. These file numbers were then used as a reference to separate the cases from the other supporting files as only reference numbers were provided in the other data files. Information containing the individual vehicle data was not originally included, so the cases were divided into two for each individual road user, and merged into the accident data using SPSS to have separate information for each vehicle within the accident. The contributory factor variables were only provided for each accident and the information was separated according to the different accident users, and merged with the main accident data file to allow for the dataset to be analysed.

7.3 Results

7.3.1 Cluster analysis

In table 41 all of the factors that were entered into the cluster analysis are outlined. A total of 18 specific variables were selected according to the most relevant risk factors that are present in multiple vehicle accidents. The age group of the road user and the gender type were included in the analysis. The first two contributory factors for each road user were included and the vehicle type was also coded. The environmental and infrastructure factors included described the road type, speed limit, junction detail, junction control and carriageway class, and were also entered into the analysis. The degree of freedom values for each included factor were included in the table. A detailed list of all of the values counts and percentages can be found in Appendix B (pp. 342).

Table 41: Stats 19 cluster analysis variables

Variable	Aspect	Level	df	Value
Speed limit	Environmental	Accident	4	≥ 30 mph; 40 mph; 50 mph; 60 mph; 70 mph
Road type	Environmental	Accident	5	Roundabout; One way; Dual carriageway; Single carriageway; Slip road; Unknown
Junction detail	Environmental	Accident	8	Roundabout; Mini roundabout; T or staggered junction; Not at junction; Slip road; Crossroads; Junction more than four arms; Private drive or entrance; Other junction
Junction control	Environmental	Accident	4	Authorised person; Automatic traffic signal; Stop sign; Give way; Uncontrolled
Light conditions	Environmental	Accident	1	Day; Night
Road user 1 & 2 gender	Road user 1 & 2	Road user	1	Male; Female
Road User 1 & 2 age group	Road user 1 & 2	Road user	5	0-17; 18-21; 21-29; 30-49; 50-65; 66+
Road user 1 & 2 contributory factors 1 & 2	Accident	Road user	9	Road environment contributed; Vehicle defects; Injudicious action; Driver/rider error or reaction; Impairment or distraction; Behaviour or inexperience; Vision affected by external factors; Pedestrian codes; Other; No factor coded
Manoeuvre	Road user 1 & 2	Road user	8	Turning let; Turning right; Waiting; Lane change; Overtaking; Going ahead left bend; Going ahead right bend; Going ahead; Other
Road type	Environmental	Accident	4	A class; B class; C class; Motorway; Minor
Road user 1 & 2 Mode of transport	Vehicle	Road user	4	Cycle; PTW; Car; LGV; HGV

The goodness of fit measurements were carried out for the two-vehicle accident cluster analysis. This analysis demonstrated that for the AIC (2096839) a 15 class solution and for the BIC (2110013) a 13 class solution was appropriate. In accordance with Linzer (2008) the BIC was used to select the number of clusters, as the number of cases was well into the thousands, and a thirteen cluster solution was selected to be analysed, the goodness of fit measures analysis can be seen in figure 21.

.

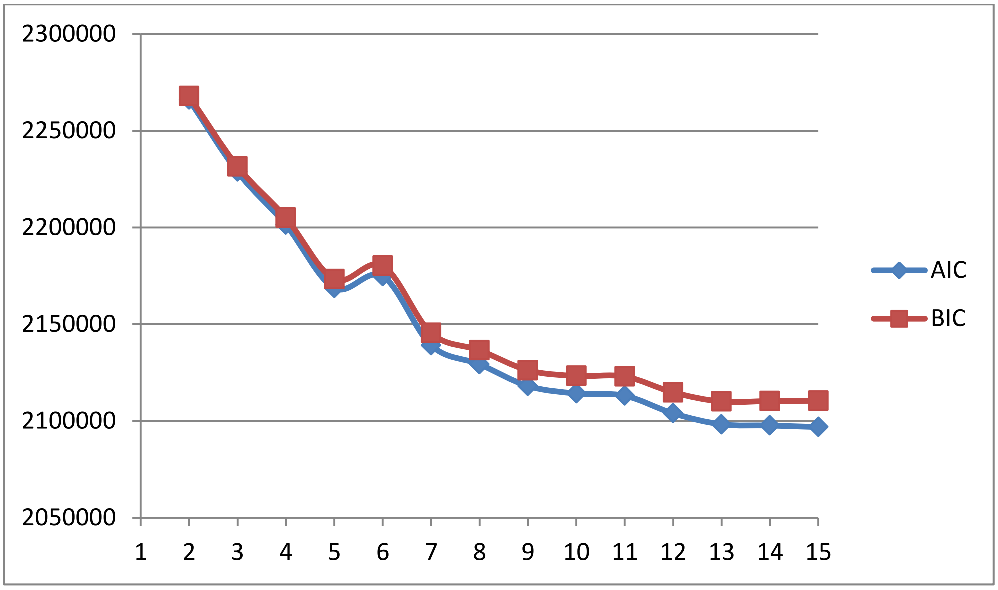


Figure 21: AIC and BIC values for Great Britain national accident data

Thirteen distinctive (separated) accident classes were highlighted resulting in a 13 solution cluster. A simple explanation of each of the clusters was carried out in order to support the discussion, these explanations can be found within the results section. The clusters were ordered with regards to case sizes. There were three large clusters that represented a third of the total cases together with 10 smaller clusters (figure 22). A detailed table of all of the cluster results can be found in Appendix B (pp. 345), in this table each over-represented significant factor is presented in bold.

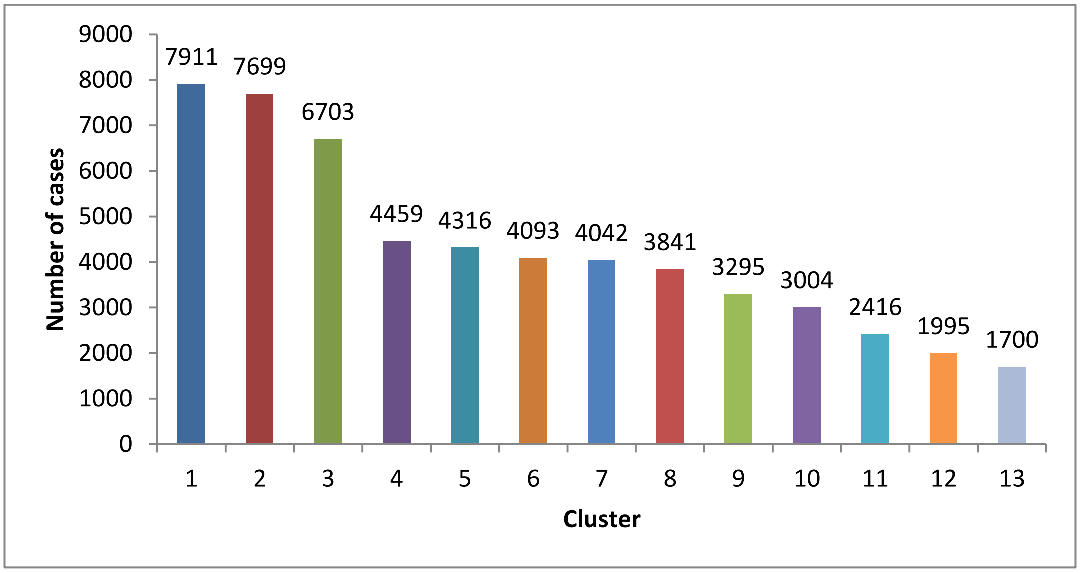


Figure 22: Two vehicle accident cluster sizes

The results from each of the clusters are presented in table form, in the below section. Due to the large number of factors that are present only factors that were significantly over-represented and accounted for at least 5% of the variance within the cluster were entered into the tables.

Cluster analysis results

Cluster 1 (n=7911)

"Accidents at a low speed give way setting while turning"

Table 42 highlights the results for cluster 1, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 42: Two vehicle cluster 1 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Female	37.8	2992	392.8	0.001
Road user 1 age group	30-49	39.9	3159	409.8	0.001
	50-65	17.9	1418		
	66+	9.6	759		
Road user 1 contributory factor 1	Error or reaction	66.2	5236	2906.6	0.001
	Vision affected by external	8.2	645		
Road user 1 contributory factor 2	Error or reaction	40.5	3200	1476.8	0.001
	Vision affected by external	7.9	623		
Road user 2 gender	Male	72.6	5744	87.5	0.001
Road user 2 age group	18-21	12.5	991	487.5	0.001
	22-29	20.1	1591		
Road user 2 contributory factor 1	No factor coded	76.0	6010	388.8	0.001
Road user 2 contributory factor 2	No factor coded	94.6	7487	499.5	0.001
Road user 1 mode of transport	Car	93.3	7383	1231.7	0.001
Road user 2 mode of transport	Cycle	18.0	1424	4624.5	0.001
	PTW	21.1	1669		
Light conditions	Day	93.7	7413	8.7	0.01
Road type	Single carriageway	94.4	7464	2373.4	0.001
Speed limit	30 mph	94.5	7473	5443.0	0.001
Junction detail	T or staggered junction	69.0	5456	7885.5	0.001
	Crossroads	15.7	1240		
	Private drive/ entrance	7.4	587		
	Other junction	5.6	439		

Junction control	Give way	93.8	7424	5708.8	0.001
Road User 1 manoeuvre	Turning left	10.9	858	1104 4.3	0.001
	Turning right	57.5	4547		
	Other	15.4	1220		
Road user 2 manoeuvre	Overtaking	5.7	453	3585. 9	0.001
	Going ahead	81.1	6415		

Human Factors

Road user 1: Female road users (37.8%) and all age groups from 30 years old and older (67.4% total) were over-represented in this cluster. The first and second contributing factor for road user 1 was error or reaction (66.2%) or vision affected by external objects (8.2%)

Road user 2: The second road user was male (72.6%) and the age groups 18-21 (12.5%) and 22-29 (20.1%) were over-represented. Both contributing factors were coded as no factor coded (76.0% & 94.6%).

Mode of transportation

The vehicle coded for road user 1 was a car (93.3%). PTWs (21.1%) and cycles (18.0%) were significantly over-represented for the second road user.

Environmental/Infrastructural Factors

The accidents occurred in a single carriageway road (94.4%) with a 30 mph or under (94.5%) speed limit at a T or staggered junction (69.0%) with a give way sign (93.8%). Road user 1 was either turning left (10.9%) or right (57.5%) and road user 2 was going ahead (81.1%).

Cluster 2 (n=7699)

“Accidents in a slow setting where the road user 1 is going ahead”

Table 43 highlights the results for cluster 2, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 43: Two vehicle cluster 2 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 age group	18-21	18.3	1407	378.2	0.001
	22-29	21.2	1630		
Road user 1 contributory factor 1	Road environment	16.9	1299	1736.1	0.001
	Injudicious action	19.1	1469		
	Impairment/Distraction	8.3	637		
	Behaviour/Inexperience	7.8	603		
Road user 1 contributory factor 2	Road environment	5.6	433	921.4	0.001
	Injudicious action	13.3	1022		
Road user 2 gender	Female	40.1	3089	275.8	0.001
Road user 2 age group	30-49	51.3	3947	269.5	0.001
	50-65	21.8	1678		
Road user 2 contributory factor 1	No factor coded	83.8	6454	949.1	0.001
Road user 2 contributory factor 2	No factor coded	97.7	519	796.3	0.001
Road user 1 mode of transport	Car	85.2	6563	444.3	0.001
Road user 2 mode of transport	Car	88.8	6834	1228.0	0.001
Light conditions	Night	8.2	632	17.0	0.001
Road type	Single carriageway	94.8	7298	2348.8	0.001
Speed limit	30 mph	62.1	4783	1007.8	0.001
	60 mph	27.6	2126		
Junction detail	T or staggered junction	58.2	4482	6070.3	0.001
	Crossroads	16.3	1254		
	Private drive/entrance	7.9	607		
	Other junction	10.0	772		
Junction control	Give way	95.2	7333	5896.6	0.001
Road User 1 manoeuvre	Going ahead left bend	8.4	644	2325.4	0.001
	Going ahead	58.3	4489		
Road user 2 manoeuvre	Turning right	8.5	657	4324.0	0.001
	Waiting	31.2	2402		
	Going ahead right bend	6.6	509		
	Other	16.8	1293		

Human Factors

Road user 1: For this cluster the age groups 18-29 (39.5% in total for two groups) were significant. The contributing factors that were over-represented were injudicious action or road factors.

Road user 2: The significant gender for road user 2 was female (40.1%). The age range was 30-49 (51.3%) or 50-65 (21.8%). The contributing factors that were significant for this road user were no factor coded.

Mode of Transportation

Road user 1 was a car driver (85.2%), and road user 2 was also a car driver (88.8%).

Environmental/Infrastructural Factors

These accidents occurred on a single carriageway (94.8%) in a 30 mph (62.1%) or 60 mph (27.6%) speed limit road in a t or staggered junction (58.2%) in a give way setting (95.2%). The first road user was going ahead (58.3%) and the second road user was waiting (31.2%) or coded as an 'other' manoeuvre (16.8%).

Cluster 3 (n=6703)

"Accidents due to injudicious actions in a single carriageway"

Table 44 highlights the results for cluster 3, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 44: Two vehicle cluster 3 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Male	76.4	5123	91.1	0.001
Road user 1 age group	0-17	10.2	682	639.9	0.001
	18-21	18.1	1210		
Road user 1 contributory factor 1	Road environment	15.5	1038	2415.0	0.001
	Injudicious action	18.2	1222		
	Impairment/Distraction	12.0	804		
	Behaviour/Inexperience	11.0	736		
Road user 1 contributory factor 2	Road environment	5.2	348	1027.1	0.001
	Injudicious action	11.4	763		
	Impairment/Distraction	5.4	363		
	Behaviour/Inexperience	11.1	746		
Road user 2 gender	Female	36.1	2421	61.5	0.001
Road user 2 age group	30-49	49.8	3336	201.1	0.001
	50-65	22.1	1478		
	66+	6.1	410		
Road user 2 contributory factor 1	No factor coded	85.1	5706	962.2	0.001
Road user 2 contributory factor 2	No factor coded	99.4	6664	931.6	0.001
Road user 1 mode of transport 1	Cycle	5.7	381	435.6	0.001
	PTW	10.9	729		
Road user 2 mode of transport	Car	86.0	5763	1019.6	0.001

	HGV	10.2	680		
Road type	Single carriageway	91.9	6157	1477.9	0.001
Speed limit	60 mph	36.6	2454	1468.7	0.001
Junction detail	No junction	100.0	6703	15535.5	0.001
Junction control	Uncontrolled	100.0	6702	15542.1	0.001
Road User 1 manoeuvre	Overtaking	9.0	604	4449.6	0.001
	Going ahead	72.6	4868		
Road user 2 manoeuvre	Waiting	14.6	975	2637.5	0.001
	Other	29.7	1987		

Human Factors

Road user 1: The significant gender for road user 1 was male (76.4%) and the age ranges between 0-21 (28.3% in total) were over-represented. The first contributing factor for this road user was injudicious action (18.2%), road environment (15.5%) or impairment/distraction (12.0%).

Road user 2: The road users gender was female (36.1%) and the three age groups above thirty years old (78.0% in total) were over-represented in this cluster. No factors coded was significant as the contributory factor for this road user

Mode of transportation

The mode of transport that was over-represented for road user 1 were cycles (5.7%) and PTWs (10.9%). The second road user was a car driver (86.0%).

Environmental/Infrastructural Factors

These accidents occurred in a single carriageway (91.9%), in a 60 mph road (36.6%) uncontrolled junction (100.0%) where the first road user was going ahead (72.6%) and the second road user was waiting (14.6%) or making a manoeuvre coded as other (29.7%).

Cluster 4 (n=4459)

“Road user error while turning”

Table 45 highlights the results for cluster 4, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 45: Two vehicle cluster 4 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Female	33.1	1475	50.0	0.001
Road user 1 age group	30-49	46.4	2068	415.1	0.001
	50-65	21.1	940		
Road user 1 contributory factor 1	No factor coded	79.9	3561	25635.0	0.001
Road user 1 contributory factor 2	No factor coded	97.5	4345	5781.8	0.001
Road user 2 gender	Male	70.9	3159	17.3	0.001
Road user 2 age group	18-21	15.1	674	764.3	0.001
	22-29	20.7	923		
	66+	7.2	322		
Road user 2 contributory factor 1	Road environment	7.1	315	13031.6	0.001
	Injudicious action	18.9	843		
	Error or reaction	59.0	2631		
	Behaviour/Inexperience	5.4	241		
Road user 1 mode of transport	PTW	8.1	361	108.7	0.001
	Car	80.8	3604		
Road user 2 mode of transport	Cycle	10.9	486	186.5	0.001
Light conditions	Day	93.2	4154	4.3	0.001
Road type	Single carriageway	79.9	3562	161.2	0.001
Speed limit	30 mph	70.2	3131	754.8	0.001
	40 mph	13.4	595		
Junction detail	Roundabout	9.6	426	2616.6	0.001
	T or staggered junction	53.3	2374		
	Crossroads	16.3	727		
	Private drive/entrance	6.8	305		
	Other junction	8.0	358		
Junction control	Traffic signal	16.1	717	2386.1	0.001
	Give way	82.3	3670		
Road User 1 manoeuvre	Waiting	16.1	718	2229.8	0.001
	Going ahead	50.4	2245		
Road user 2 manoeuvre	Turning left	5.1	227	3104.4	0.001
	Turning right	25.9	1156		
	Overtaking	5.1	226		

Human Factors

Road user 1: The gender of the first road user was female (33.1%) and the age ranges between 30-65 (67.5%) were over-represented. The contributing factors were both no factor coded.

Road user 2: Males (70.9%) and the age range between 18-29 (35.8%) and 66 years or older (7.2%) was over-represented. The contributing risk factor was coded as error or reaction (59.0%).

Mode of transportation

The vehicle coded for road user 1 was a car (80.8%). Cycles (10.9%) were over-represented for the second road user.

Environmental/Infrastructural Factors

The accidents occurred on a single carriageway road (79.9%) during the day (93.2%) at a T junction (53.3%) and a give way setting (82.3%). The first road user was going ahead (50.4%) and the second road user was turning right (25.9%).

Cluster 5 (n=4316)

"Intersection accidents due to breaking the law"

Table 46 highlights the results for cluster 5, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 46: Two vehicle cluster 5 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Female	36.6	1578	149.3	0.001
Road user 1 age group	50-65	20.5	886	1028.5	0.001
	66+	18.5	798		
Road user 1 contributory factor 1	Error or reaction	71.2	3074	1613.9	0.001
	Vision affected by external	5.2	223		
Road user 1 contributory factor 2	Error or reaction	52.4	2263	1330.1	0.001
	Vision affected by external	5.7	244		
Road user 2 gender	Male	71.4	3081	23.6	0.001
Road user 2 age group	18-21	10.0	429	14.6	0.001
	22-29	17.4	749		
	50-65	20.5	883		
	66+	5.7	246		
Road user 2 contributory factor 1	Injudicious action	5.2	224	364.1	0.001
	No factor coded	80.4	3469		
Road user 2 contributory factor 2	No factor coded	96.2	4153	313.3	0.001
Road user 1 mode of transport	Car	89.5	3862	428.8	0.001
Road user 2 mode of transport	PTW	17.1	737	576.6	0.001
Light conditions	Day	93.4	4029	4.3	0.001
Road type	Single carriageway	86.9	3750	600.6	0.001

Speed limit	40 mph	19.7	852	2862.5	0.001
	50 mph	5.7	247		
	60 mph	46.3	1999		
Junction detail	T or staggered junction	68.8	2968	4434.3	0.001
	Private drive/entrance	11.8	511		
	Other junction	6.1	262		
Junction control	Stop sign	2.5	106	3380.8	0.001
	Give way	96.4	4158		
Road User 1 manoeuvre	Turning left	6.1	263	9569.2	0.001
	Turning right	75.4	3252		
Road user 2 manoeuvre	Overtaking	4.7	203	2429.0	0.001
	Going ahead	85.9	3706		

Human Factors

Road user 1: Female (36.6%) road users and the age groups 50-65 (20.5%) and 66+ (18.5%) were over-represented for this cluster. The contributing factor error or reaction was over-represented for both contributing factors.

Road user 2: Road user 2 was coded as a male (71.4%) and all age ranges other than 0-17 and 30-49 were over-represented for this cluster. The contributing factor was no factor coded.

Mode of transportation

The vehicle coded for road user 1 was a car (89.5%). PTWs (17.1%) were over-represented for the second road user.

Environmental/Infrastructural Factors

The accidents occurred in a single carriageway setting (86.9%) during the day in a 40-60 mph speed limit, at a T junction (68.8%) at a give way sign (96.4%). The first road user was turning right (75.4%) and the second road user was going ahead (85.9%).

Cluster 6 (n=4093)

"Road user incorrectly entering a roundabout"

Table 47 highlights the results for cluster 6, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 47: Two vehicle cluster 6 analysis results

Variable	Value	Percentage	N	χ²	Sig.
Road user 1 gender	Female	30.4	1245	8.0	0.01
Road user 1 age group	30-49	40.5	1656	94.8	0.001
	50-65	18.7	764		
	66+	8.6	352		
Road user 1 contributory factor 1	Injudicious action	14.9	609	553.8	0.001
	Error or reaction	59.1	2417		
	Pedestrian only	2.9	119		
Road user 1 contributory factor 2	Error or reaction	35.1	1438	106.4	0.001
Road user 2 gender	Female	34.5	1412	13.4	0.001
Road user 2 contributory factor 1	Error or reaction	14.1	578	144.5	0.001
	Impairment/Distraction	1.6	64		
	No factor coded	74.7	3058		
Road user 2 contributory factor 2	No factor coded	93.6	3829	141.7	0.001
Road user 1 mode of transport	Car	80.6	3299	31.1	0.001
	HGV	10.2	417		
Road user 2 mode of transport	Cycle	13.3	545	486.6	0.001
	PTW	12.1	495		
Light conditions	Day	93.2	3813	4.3	0.05
Road type	Roundabout	92.4	3783	44459.2	0.001
Speed limit	30 mph	57.3	2343	292.6	0.001
	40 mph	16.3	669		
	50 mph	3.9	158		
Junction detail	Roundabout	91.3	3739	38476.0	0.001
	Mini roundabout	8.5	347		
Junction control	Authorised person	0.3	12	2543.7	0.001
	Give way	92.1	3770		
Road User 1 manoeuvre	Turning left	7.1	289	958.6	0.001
	Waiting	5.3	217		
	Going ahead	47.1	1928		
	Other	22.3	912		
Road user 2 manoeuvre	Turning right	9.9	404	455.2	0.001
	Waiting	17.0	697		

Human Factors

Vehicle 1: The first road users gender was female (30.4%) and age range was 30 years old and older (67.8%). The contributing factors for road user 1 were error or reaction.

Vehicle 2: The second road user was female (34.5%) and age range related variables were not significant. The first contributing factor for road user 2 was no factor coded (74.7%).

Mode of transportation

The vehicle coded for road user 1 was a car (80.6%). Cycles (13.3%) or PTWs (12.1%) were over-represented for the second road user.

Environmental/Infrastructural Factors

The accidents predominantly occurred during the day (93.2%) on a roundabout (92.4%) in a 30 mph (57.3%) speed limit road in a give way setting (92.1%). Road user 1 was going ahead (47.1%) and road user 2 was waiting (17.0%).

Cluster 7 (n=4042)

"Road user error at traffic lights"

Table 48 highlights the results for cluster 7, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 48: Two vehicle cluster 7 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 age group 2	18-21	16.3	660	116.9	0.001
	22-29	22.4	907		
Road user 1 contributory factor 1	Injudicious action	27.8	1122	1069.2	0.001
Road user 1 contributory factor 2	Injudicious action	8.5	343	236.6	0.001
	Error or reaction	31.9	1288		
	Behaviour/Inexperience	9.5	385		
Road user 2 age group	22-29	20.3	822	62.7	0.001
Road user 2 contributory factor 1	Injudicious action	6.5	262	332.3	0.001
	No factor coded	77.8	3145		
Road user 2 contributory factor 2	No factor coded	94.4	3817	230.1	0.001
Road user 1 mode of transport	Car	85.1	3441	172.8	0.001
	LGV	2.7	107		
Road user 2 mode of transport	Car	82.9	3350	204.1	0.001
	LGV	3.0	123		
Road type	Dual carriageway	36.6	1480	1689.9	0.001
Speed limit	30 mph	76.6	3097	1602.1	0.001
	40 mph	17.1	692		
Junction detail	Crossroads	59.8	2416	13138.7	0.001
	Four or more arms	11.2	453		

Junction control	Traffic signal	89.6	3620	20972.6	0.001
Road User 1 manoeuvre	Turning right	34.6	1398	1118.0	0.001
	Going ahead	45.4	1836		
Road user 2 manoeuvre	Turning right	11.3	455	570.1	0.001
	Waiting	12.9	521		
	Going ahead	60.8	2457		

Human Factors

Road user 1: The age range 18-29 (38.7%) was over-represented. The contributory factors injudicious action and error or reaction were over-represented for this road user.

Road user 2: The age range 22-29 (20.3%) was significant for this road user. The contributory factors no factor coded were significant.

Mode of transportation

The first road user was a car (85.1%). The vehicle type for the second road user was also a car (82.9%).

Environmental/Infrastructural Factors

The main factors that were outlined in this analysis were that the accidents occurred on a 30 mph (76.6%) speed limit dual carriageway road (36.6%) at a cross roads (59.8%) with a traffic signal control (89.6%). Road user 1 was going ahead (45.4%) or turning right (34.6%) and road user 2 was going ahead (60.8%).

Cluster 8 (n=3841)

"High speed accidents on a motorway due to faulty manoeuvre"

Table 49 highlights the results for cluster 8, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 49: Two vehicle cluster 8 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Male	78.2	3003	90.7	0.001
Road user 1 age group	22-29	21.0	806	342.6	0.001
	30-49	46.9	1801		
	50-65	17.3	664		
Road user 1 contributory factor 1	Error or reaction	45.8	1760	188.6	0.001
	Impairment/Distraction	7.3	280		
	Vision affected by external	6.1	234		
Road user 1 contributory factor 2	Injudicious action	8.3	320	105.1	0.001
	No factor coded	45.7	1754		
Road user 2 gender	Male	72.4	2782	36.3	0.001
Road user 2 age group	30-49	51.0	1960	145.6	0.001
	50-65	22.0	844		
Road user 2 contributory factor 1	No factor coded	74.4	2859	69.0	0.001
Road user 2 contributory factor 2	No factor coded	92.3	3546	75.3	0.001
Road user 1 mode of transport	HGV	29.1	1119	2108.6	0.001
Road user 2 mode of transport	Car	76.0	2917	1101.1	0.001
	HGV	19.6	752		
Road type	Dual carriageway	97.2	3733	20539.9	0.001
	Slip road	2.0	77		
Speed limit	50 mph	5.6	215	31831.0	0.001
	70 mph	86.7	3329		
Junction detail	No junction	100.0	3841	8406.1	0.001
Junction control	Uncontrolled	100.0	3841	8414.7	0.001
Road User 1 manoeuvre	Lane change	24.0	923	6998.6	0.001
	Overtaking	7.3	279		
	Going ahead	51.9	1992		
Road user 2 manoeuvre	Lane change	5.2	198	1398.2	0.001
	Going ahead	59.8	2297		
	Other	18.7	717		

Human Factors

Road User 1: The gender of the first road user was male (78.2%) and the age ranges between 22-65 (88.2%) were significant for this cluster. The first contributing factor was error or reaction (45.8%).

Vehicle 2: The gender of the second road user was male (72.4%) and the age ranges between 30-65 (73.0%) were significant. The contributing factor was no factor coded.

Vehicular Factors

The vehicle coded for road user 1 was a HGV (29.1%). The second road user was either a car (76.0%) or a HGV (19.8%) driver.

Environmental/Infrastructural Factors

The accidents occurred on a dual carriageway (97.2%), in a 70 mph speed limit (86.7%) road with no junction (100.0%) or traffic control (100.0%). Both road users were going ahead or road user 1 was making a lane change or overtaking manoeuvre.

Cluster 9 (n=3295)

"Road environment or behaviour errors on a bend"

Table 50 highlights the results for cluster 9, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 50: Two vehicle cluster 9 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Male	77.0	2537	52.1	0.001
Road user 1 age group	0-17	5.9	194	210.1	0.001
	18-21	19.1	629		
	22-29	22.9	756		
Road user 1 contributory factor 1	Road environment	37.7	1243	3492.7	0.001
	Injudicious action	15.7	518		
	Behaviour/Inexperience	9.7	319		
Road user 1 contributory factor 2	Road environment	16.3	537	2581.7	0.001
	Injudicious action	14.7	485		
	Behaviour/Inexperience	9.9	327		
	Vision affected by external	4.9	161		
Road user 2 age group	30-49	48.2	1587	200.2	0.001
	50-65	25.9	854		
	66+	6.8	222		
Road user 2 contributory factor 1	Road environment	9.3	307	807.9	0.001
	Vision affected by external	4.3	141		
	No factor coded	75.4	2486		
Road user 2 contributory factor 2	Road environment	2.7	89	450.1	0.001
	Vision affected by external	1.6	52		
	No factor coded	92.7	3055		
Road user 1 mode of transport	PTW	9.3	307	99.4	0.001

	Car	80.1	2639		
Road user 2 mode of transport	Car	80.3	2645	485.2	0.001
	HGV	13.7	451		
Light conditions	Night	11.0	362	81.3	0.001
Road type	Single carriageway	97.0	3196	1047.6	0.001
Speed limit	60 mph	71.2	2346	5040.0	0.001
Junction detail	No junction	100.0	3295	7140.5	0.001
Junction control	Uncontrolled	100.0	3295	7147.8	0.001
	Unknown	0.8	28		
Road User 1 manoeuvre	Going ahead left bend	60.6	1997	28713/7	0.001
	Going ahead right bend	30.4	1000		
Road user 2 manoeuvre	Going ahead left bend	26.4	869	21295.8	0.001
	Going ahead right bend	50.9	1677		

Human Factors

Vehicle 1: The first road users gender was male (77.0%) and the age range 18-29 (42.0%) was over-represented. The first contributing factors for road user 1 were road environment or injudicious action.

Vehicle 2: The second road user did not have an over-represented gender. The age range 30 years old or older was over-represented (80.9%). The contributing factor was no factor coded.

Mode of transportation

The vehicle coded for both road users were cars.

Environmental/Infrastructural Factors

The accidents predominantly occurred on a single carriageway (97.0) in an uncontrolled junction (100.0%) in a 60 mph road (71.2%) on a bend**.**

Cluster 10 (n=3004)

"Vulnerable road user fails to give way"

Table 51 highlights the results for cluster 10, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 51: Two vehicle cluster 10 analysis results

Variable	Value	Percentage	N	χ²	Sig.
Road user 1 gender	Male	92.1	2767	663.5	0.001
Road user 1 age group	0-17	38.7	1163	7791.2	0.001
	18-21	18.1	542		
Road user 1 contributory factor 1	Injudicious action	23.3	698	947.5	0.001
	Behaviour/Inexperience	13.9	418		
Road user 1 contributory factor 2	Injudicious action	9.4	283	527.6	0.001
	Error or reaction	30.2	907		
	Behaviour/Inexperience	15.8	476		
Road user 2 gender	Female	35.8	1074	21.3	0.001
Road user 2 age group	22-29	18.0	541	41.6	0.001
	50-65	20.5	615		
	66+	5.9	177		
Road user 2 contributory factor 1	Error or reaction	15.6	470	253.4	0.001
	No factor coded	72.5	2179		
Road user 2 contributory factor 2	No factor coded	92.5	2778	136.6	0.001
Road user 1 mode of transport	Cycle	26.6	799	13600.5	0.001
	PTW	43.7	1314	450.4	0.001
Road user 2 mode of transport	Car	89.8	2697		
Light conditions	Day	97.2	2919	87.1	0.001
Road type	Single carriageway	92.6	2783	716.0	0.001
Speed limit	30 mph	84.4	2534	1049.4	0.001
Junction detail	T or staggered junction	65.5	1969	2747.1	0.001
	Private drive/entrance	11.5	345		
	Other junction	7.6	230		
Junction control	Give way	94.4	2837	2007.7	0.001
Road User 1 manoeuvre	Overtaking	22.6	678	2850.2	0.001
	Going ahead	56.5	1697		
Road User 2 manoeuvre	Turning left	5.7	170	3479.2	0.001
	Turning right	33.0	992		

Human Factors

Vehicle 1: The first road users gender was male (92.1%) and age range was 0-21 (56.8%). The first contributing factors were injudicious action and error/reaction.

Vehicle 2: The second road users gender was female (35.8%) and the age ranges 22-29 (18.0%) and 50 and above (26.4%) was over-represented. The contributory factor was no factor coded.

Mode of transportation

The vehicle coded for road user 1 was a cycle (26.6%) or PTW (43.7%). The second road user was a car driver (89.8%).

Environmental/Infrastructural Factors

The accidents occurred in a single carriageway (92.6%) at a T junction (65.5%) at a give way sign (94.4%). Road user 1 was going ahead (56.5%) and road user 2 was turning left (5.7%) or right (33.0%).

Cluster 11 (n=2416)

"Younger road user accident on a bend"

Table 52 highlights the results for cluster 11, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 52: Two vehicle cluster 11 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Male	77.0	1860	52.1	0.001
Road user 1 age group	0-17	5.9	142	210.1	0.001
	18-21	19.1	461		
	22-29	22.9	554		
Road user 1 contributory factor 1	Road environment	37.7	912	3492.7	0.001
	Injudicious action	15.7	380		
	Behaviour/Inexperience	9.7	234		
Road user 1 contributory factor 2	Road environment	16.3	394	2581.7	0.001
	Injudicious action	14.7	356		
	Behaviour/Inexperience	9.9	240		
	Vision affected by external	4.9	118		
Road user 2 age group	30-49	48.2	1164	200.2	0.001
	50-65	25.9	626		
	66+	6.8	163		
Road user 2 contributory factor 1	Road environment	9.3	225	807.9	0.001
	No factor coded	75.4	1823		
Road user 2 contributory factor 2	No factor coded	92.7	2240	450.1	0.001
Road user 1 mode of transport 2	PTW	9.3	225	99.4	0.001
	Car	80.1	1935		
Road user 2 mode of transport	Car	80.3	1939	485.2	0.001
	HGV	13.7	331		
Light conditions	Night	11.0	266	81.3	0.001

Road type	Single carriageway	97.0	2343	1047.6	0.001
Speed limit	60 mph	71.2	1720	5040.0	0.001
Junction detail	No junction	100.0	2416	7140.5	0.001
Junction control	Uncontrolled	100.0	2416	7147.8	0.001
Road User 1 manoeuvre	Going ahead left bend	60.6	1465	28713.7	0.001
	Going ahead right bend	30.4	733		
Road user 2 manoeuvre	Going ahead left bend	26.4	637	21295.8	0.001
	Going ahead right bend	50.9	1230		

Human Factors

Vehicle 1: The first road users gender was male (77.0%) and age range was between 0-29 (47.9% in total). The contributing factors were road environment or injudicious action.

Vehicle 2: The road user 2 age range 30 or older (80.9%) were significant. The first contributing factor was no factor coded (75.4%).

Mode of transportation

The vehicle coded for road user 1 was a car (80.1%) and the second road user was coded as a car (80.3%).

Environmental/Infrastructural Factors

The accidents predominantly occurred in a single carriageway (97.0%) road with no junction (100.0%) on a bend (91.0%).

Cluster 12 (n=1995)

“High speed accidents on a motorway due to faulty manoeuvre”

Table 53 highlights the results for cluster 12, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 53: Two vehicle cluster 12 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Female	30.5	609	4.2	0.05
Road user 1 age group	30-49	44.5	888	179.0	0.001

	50-65	22.7	452		
Road user 1 contributory factor 1	No factor coded	65.6	1309	6984.1	0.001
Road user 1 contributory factor 2	No factor coded	84.7	1689	1465.7	0.001
Road user 2 gender	Male	73.0	1457	23.3	0.001
Road user 2 age group	18-21	17.4	347	426.9	0.001
	22-29	18.9	376		
Road user 2 contributory factor 1	Road environment	17.0	339	5243.6	0.001
	Injudicious action	16.2	324		
	Error or reaction	44.7	891		
	Impairment/Distraction	5.4	108		
	Behaviour/Inexperience	8.4	167		
Road user 2 contributory factor 2	Injudicious action	10.2	204	6778.6	0.001
	Error or reaction	32.0	638		
	Behaviour/Inexperience	9.7	193		
Road user 1 mode of transport	Car	80.0	1596	39.4	0.001
Road user 2 mode of transport	Cycle	10.0	199	700.8	0.001
	PTW	12.2	243		
Light conditions	Night	8.9	177	9.7	0.01
Road type	Single carriageway	87.6	1748	288.8	0.001
Speed limit	40 mph	11.5	230	222.8	0.001
	60 mph	31.0	619		
Junction detail	No junction	100.0	1995	4222.5	0.001
Junction control	Uncontrolled	100.0	1994	4220.5	0.001
Road User 1 manoeuvre	Waiting	10.4	206	913.5	0.001
	Overtaking	5.8	116		
	Going ahead right bend	8.7	173		
	Going ahead	42.7	851		
	Other	24.4	486		
Road user 2 manoeuvre	Overtaking	8.1	161	700.8	0.001
	Going ahead left bend	10.0	199		
	Other	17.4	348		

Human Factors

Vehicle 1: The first road users gender was female (30.5%) and age range was between 30-65 (67.2%). The contributing factors were no factor coded.

Vehicle 2: The second road users gender was male (73.0%) and age range was between 18-29 (36.3%). The contributory factors were error or reaction and road environment.

Mode of transportation

The vehicle coded for road user 1 was a car (80.0%) and the second road user's mode of transportation was over-represented as a cycle (10.0%) or a PTW (12.2%).

Environmental/Infrastructural Factors

The accidents occurred on s single carriageway road (87.6%), where road user 1 was going ahead (43.7%) and road user 2 was making an 'other' manoeuvre (17.4%).

Cluster 13 (n=1700)

"Junction accident due to error"

Table 54 highlights the results for cluster 13, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 54: Two vehicle cluster 13 analysis results

Variable	Value	Percentage	N	χ^2	Sig.
Road user 1 gender	Male	73.8	1255	44.7	0.05
Road user 1 age group	22-29	21.8	371	72.3	0.001
	30-49	42.1	716		
	50-65	17.5	297		
Road user 1 contributory factor 1	Error or reaction	53.0	901	81.1	0.001
	No factor coded	47.6	809		
Road user 2 age group	30-49	50.0	850	38.7	0.001
	50-65	21.6	366		
Road user 2 contributory factor 1	Error or reaction	15.3	260	49.3	0.001
	No factor coded	73.2	1244		
Road user 2 contributory factor 2	No factor coded	91.5	1555	32.1	0.001
Road user 1 mode of transport	HGV	17.2	293	201.1	0.001
	Car	81.1	1378		
Road user 2 mode of transport	HGV	12.6	215	175.2	0.001
Road type	Dual carriageway	69.6	1183	10913.6	0.001
	Slip road	20.8	353		
Speed limit	40 mph	11.5	195	7259.7	0.001
	50 mph	8.0	136		
	70 mph	64.5	1096		
Junction detail	Roundabout	37.0	629	17008.7	0.001
	Slip road	39.1	665		

	Other junction	7.5	127		
Junction control	Give way	89.0	1514	951.6	0.001
Road User 1 manoeuvre	Waiting	7.7	131	1255.8	0.001
	Lane change	15.5	263		
	Going ahead	42.7	725		
	Other	21.4	363		
Road user 2 manoeuvre	Waiting	22.4	381	572.3	0.001
	Other	18.7	318		

Human Factors

Vehicle 1: The first road users gender was male (73.8%) and age range was between 22-65 (81.4%). The contributing factors for road user 1 were error or reaction (53.0%).

Vehicle 2: The second road user did not have a significant value for gender and the age ranges 30-65 were over-represented (71.6%). The contributing factors were no factor coded.

Mode of transportation

The vehicle coded for road user 1 was a car (81.1%) and the second road user was coded as a HGV (12.6%).

Environmental/Infrastructural Factors

The accidents occurred on a dual carriageway (69.6%) on a roundabout (37.0%) or slip road (39.1%) with a give way (89.0%) sign.

7.4 Discussion

7.4.1 National data compared to in-depth data using cluster analysis

There is a need to take a few considerations into account before a detailed understanding of the results from the cluster analysis can be made. First of all it is necessary to consider the differences between the contributory factors reporting scheme that was used for the STATS19 coding with the more detailed HFF accident coding method.

All contributory factors were included in this analysis, though for a large number of the cases only one road user was coded as having a contributory factor. So a direct understanding of what both road users did was sometimes not possible. In 79.7% of the accidents a contributory factor was only attributed to one of the road users. This demonstrates that there are limitations in terms of coding accident causation behaviour particularly when the accidents are reported to the police officers rather than collected on scene.

The more detailed level of information in terms of the accident site variables skewed the clusters towards coding these variables, and without the contributory factor information the exact understanding of the failures was not possible. The nature of the STATS19 data did not allow for groupings of the accident cases based on accident type, so the main grouping possible was based on the carriageway characteristics of the accidents.

When interpreting the cluster analysis results it can be seen that 4 of the 6 largest clusters were related to single carriageway accidents on roads that had a T junction. The contributory factor codes did not allow for a detailed analysis of the accident causation portion of these accidents, though possible coding with the vehicle point of impact may have provided more detailed information with regards to the accident type. Due to these issues the clustering algorithm turned slight differences for this accident types into different cluster groupings.

A way around these issues may be to only select one manoeuvre for analysis or to only use select police reports that provide detailed information about specific cases and use these as a sampling tool such as Clarke et al. (1999). When using specific cases the data is reprocessed into a new database to allow for causal inferences to be developed, rather than using raw data provided by databases such as STATS19.

A further issue is the difference in the detail level of the cases collected. For the STATS19 data a total of 82 fields can be possibly completed and a number of the cases are collected retrospectively with a large amount of the data not being collected.

The OTS data collected could gather up to 3,000 variables and as the cases are collected on the spot the cases are fundamentally more complete compared to the STATS19 cases.

7.4.2 Issues when analysing accident data

A number of statistical methods have been used to analyse accident data, the use of these methods relies upon the research question that the researchers is asking. Some of the fundamental characteristics of accident data result in methodological limitations that are not fully understood (Savolainen, Mannering, Lord, & Quddus, 2011).

Some of these limitations have been identified as the underreporting of accidents, the ordinal nature of injury data, omitted variable bias, the difficulty in capturing behaviour related factors, ignoring factors related to space and time, and small sample size (Lord & Mannering, 2010; Mannering & Bhat, 2014; Savolainen et al., 2011).

Lord and Mannering (2010) identified 13 different types of data modelling tools that are currently used to analyse accident data. The accident data used for analysis purposes for research is most commonly either national or regional data. The nature of bivariate and multivariate analysis tools requires certain parameters and estimations to be made and set for the data analysis to be plausible and meaningful.

When analysing accident data, a number of considerations need to be made depending on the type of statistical analysis that is used and interpretation outcomes that is aimed for. Some of the main points of consideration are;

1. Completeness of the data
2. Exposure
3. Reported accidents

In order for the analysis of the data to be possible it is necessary that the data be as complete as possible. In cases where there are missing values a number of possible methods are available to use to replace the data. If the fact that data are missing does not depend upon any values, or potential values, for any of the variables, then data are said to be ‘missing completely at random’. This is an ideal situation for missing data as any observation is just as likely as another to be missing.

If the data available can predict the unavailable data than a ‘missing at random’ computation can be used to predict the remaining data (Hautzinger et al., 2007). This computation estimates the values of the missing data by including a number of variables that are relevant to this information. For example if we are considering whether the vehicles running lights were on or off then it would be necessary to also include the variable day and night and also weather conditions into the analysis to allow for a probable calculation (Hautzinger et al., 2007).

A third type of data is the ‘missing not at random’ data type. This situation occurs when the data is not even missing at random and so a model that describes these occurrences is required to be developed. The underlying missing factors are important and difficult to decipher.

Exposure methods aim to estimate the relative and absolute crash risks of different road user types (Huang et al., 2011). The number of trips that people make and the number of times that a risk is possible according to these trips. It is difficult to find exposure data for accident safety research, as the number of variables included in these datasets is quite large and definitive and suitable data for this source is not possible.

Reported accidents refer to whether an accident is reported or not to the police. When considering non-fatal accidents the under reporting of accidents is quite well known (Amoros, Martin, & Laumon, 2003). The degree of under reporting can be quite large and needs to be addressed by each country specifically for accident analysis purposes.

7.5 Summary

Chapter 7 presented an analysis of Great Britain national data using the relevant factor coding sheets to compare these coding against the multiple vehicle cluster analysis carried out in chapter 6.

This analysis demonstrated that when a detailed level of information is not included in the cluster analysis, the results tend to be skewed towards physical variables that provide a greater amount of information.

A discussion with regards to the differences between in-depth and national data analysis procedures and finding implications was carried out. Issues to consider when using accident data were highlighted with regards to the analysis.

8 An Analysis of Powered Two Wheeler accidents

8.1 Introduction

The literature review carried out in section 2.14.1 established that powered two wheeler (PTW) riders are one of the most at risk road user groups within the traffic environment. Statistical data shows that each year they represent 15% of people killed on European roads, and according to the World Health Organization nearly 200,000 deaths in the world annually (WHO, 2006).

Though the literature (Clabaux et al., 2012; Clarke et al., 2007; Haque et al., 2009; MAIDS, 2009) provides detailed information about PTW accidents, an approach using each case to observe the riders and other road users contribution to each accident, and defining each case according to the failure that each road user makes, would benefit in providing PTW scenarios where the interactions of both road users would be quantified against each other. This study was carried out to focus on how PTW accidents occur on an accident basis, and to identify the two way relationships between riders and other road users in accidents involving PTW riders and other types of vehicles.

A retrospective analysis using accident causation coding for each PTW accident occurring in the OTS dataset between the years 2000-2010 was carried out. Four hundred and forty nine cases were coded and analysed in terms of relevant factors to understand how PTW accidents occurred.

The aim of this study was to understand the different type of failures that PTW riders make when involved in either a single vehicle accident or when interacting on the roadway with other vehicles in a multiple vehicle accident. This study aimed to distinguish the factors and situations that were found in different PTW accident groupings. Analysis was conducted in two steps, first using descriptive methods to understand the data and then performing a latent class cluster analysis to group similar accidents together in order to identify accident scenarios.

8.2 Method

8.2.1 Design

This study uses a large number of in-depth accidents collected on site as part of the On the Spot accident study (OTS) carried out in the UK from 2000 to 2010. Factors relating to the accident were obtained by grouping the accident variables into 4 specific groups relating to human, vehicular, infrastructural or environmental factors relating to the accident. The larger number of years included compared to the study carried out in chapter 6 was a result of the necessary number of cases for the cluster analysis. As the number of cases involving PTWs between 2000-2003 were 147 total cases and were not sufficient for a detailed cluster analysis to be carried out, all of the cases collected within the OTS study needed to be included.

8.2.2 Sample

Accident data was collected by two separate groups in two different areas. The results presented in this study are based on 4,004 accidents involving a total of 12,749 vehicles and 527 pedestrians. From these accidents 449 accidents involving Powered Two-Wheelers were selected to be analysed. Of these cases 21 did not have sufficient data to be included in the cluster analysis and were thus excluded. The total number of PTW accidents included in the cluster analysis was 428.

In the sample of the PTW riders 400 of the individuals were male and 45 were female while 4 of the individuals’ gender was unknown. Of the 429 individuals whose age was coded in the sample the age for the riders were on average 32.2 years old with a standard deviation of 13.5 (figure 23).

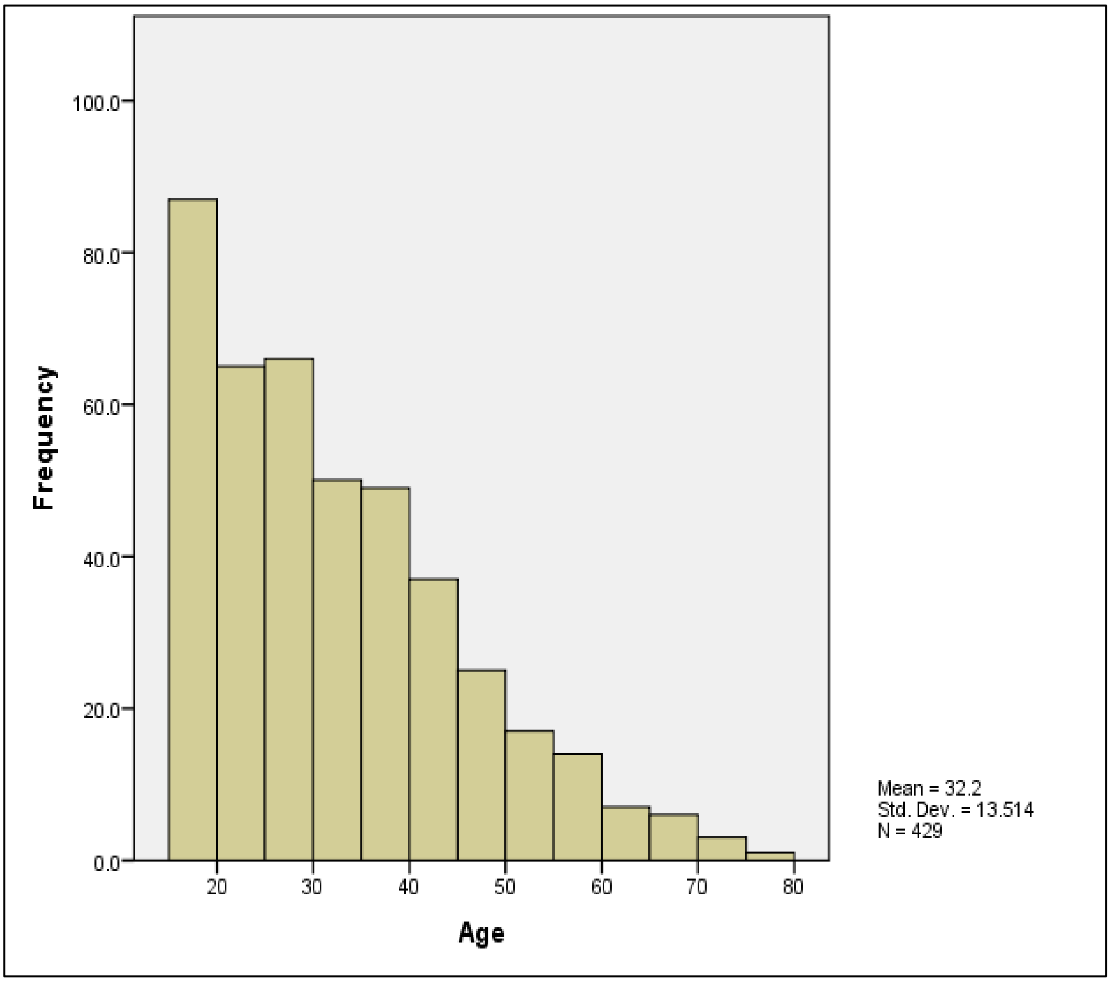


Figure 23: PTW rider age distribution

8.2.3 Procedure

For this study all of the PTW accidents present in the OTS mass data file that contained relevant data in terms of the vehicle, environment, infrastructure and human participant in relation to the accident was selected and separated into a new data file. All of the cases were retrospectively analysed by the author using the Human Functional Failure Causation methodology (Naing et al., 2007) and the LAB accident type coding diagrams by deducing and reporting the causal sequence related to each PTW accident and identifying the typical accident image that occurred. This data file was then merged with the OTS data file.

The analysis was carried out in two steps. First a descriptive analysis was carried out on the merged data to identify specific accident characteristics and for data interpretation purposes. The second stage of analysis was a latent class cluster analysis. This analysis was carried out by identifying variables that are important in the literature, and by using the descriptive analysis and literature as a guide, to separate variables into groupings that would be meaningful. The variables were grouped in a larger higher level group so the cluster analysis would not discount them. For example, variables such as in a hurry, panic and the road user's emotional status were grouped into one variable termed physical/physiological. These groupings were based on the HFF method variable groups that were described in chapter 4.

8.2.4 Statistical analysis

This study incorporated a cluster analysis in order to group the accidents in several collision scenarios. Accident causation data fields are categorical data in nature, so it was necessary to handle this data in an appropriate manner. The handling procedure was carried out by separating the specific factors and entering them into a latent class cluster analysis.

The factors that contributed to the accident and analysed in the cluster analysis are illustrated in Table 61. In addition to these variables the road type and environment variables were selected from the recorded data by the accident investigators, these factors can also be seen in Table 61. Due to the nature of the analysis it was necessary to identify the main failure factors that are most prominently present in both single and multiple vehicle PTW accidents. All PTW accidents were included in the cluster analysis whether single or multi-vehicle accidents. This was done by using the HFF chain analysis method.

The latent class cluster analysis was run including all of the above described factors above fitting them on between 2 to 15 clusters. From these clusters a comparison of the AIC and BIC was made to optimise goodness of fit while making sure that the degrees of freedom of the analysis was not negative. A chi-square analysis for all of the individual factors in each individual cluster was conducted against the total value for the dataset. So the analysis that was carried out was an analysis of the first cluster versus the total and expected values, the second cluster against the total and expected values and so forth.

For each variable that was seen to be significant a residual analysis of the values were then conducted to further determine the value that was statistically significant in the cluster analysis. Each residual value that was above 2 and also higher than the expected value was entered into a table for that specific cluster and described in the description section for each specific cluster.

8.3 Results

8.3.1 Descriptive analysis

Below is a description of the different types of PTW accidents that were analysed within this study. Table 55 presents the number of vehicle records for all of the accidents that are present within the dataset. The first road user within the accident was always coded as a PTW, and the interacting vehicles (n=340) other road users to PTW accidents were coded. In total 449 accidents were coded. Of these accidents 109 were single vehicle accidents, 302 were two vehicle accidents, 30 were three vehicle accidents, and 8 were four vehicle accidents. The other road user most commonly interacting with the PTW riders in multiple vehicle accidents were car (87.9%) drivers and for this reason the interacting road user type was not included in the cluster analysis, rather the failure mechanism that the other road user made was included.

Table 56 demonstrates all of the different types of failures for multiple vehicle accidents with the first combination of rider and road user that interacted within the accident. This was done by identifying the failure types and comparing rider and road user types. The coding illustrates that for detection failures made by the interacting road user a majority of the cases were prognosis failures (n=156), where the PTW user did not expect the driver to make a manoeuvre. The same accident type was also identified for cases for where the road user was not identified by the rider (n=41).

Table 55: PTW rider by other road user type and number of vehicles involved in the accident

	PTW accidents				
Interacting Vehicle	Single Vehicle	Two Vehicle	Three Vehicle	Four Vehicle	N
Car	0	275	18	6	299
Heavy Goods	0	4	1	0	5
Light Goods	0	14	1	1	16
PTW	109	0	2	0	111
Pedestrian	0	7	4	1	12
Other	0	10	4	0	14
Total	109	302	30	8	449

This is the largest group of failures and includes both ‘looked but did not see’ accidents or ‘right of way violation’ accident types. A large number of accidents occurred that were road users making decision errors and undertaking a risky driving behaviour and colliding with a PTW rider (n=40). Other failure types that had a relevant number of cases were diagnosis failures of the roadway by the PTW rider (n=29) and decision failures (risky behaviours undertaken) made by the PTW rider (n=36).

Table 56: Failure types for PTW riders compared to interacting road users

PTW	Interacting road user failure					
Failure	Detection	Diagnosis	Prognosis	Decision	Other	N
Detection	7	1	41	4	1	54
Diagnosis	3	1	24	1	0	29
Prognosis	156	13	5	28	5	207
Decision	11	1	18	6	0	36
Execution	1	1	3	1	0	6
Overall	0	0	7	0	0	7
Total	178	17	81	40	6	339

Table 57 demonstrates a cross tabulation between the different failure types by a number of demographic variables, contributory factors, PTW size, level of accident involvement, environmental factors and injury severity. All factors could be specifically coded for all of the accidents and so the totals do not necessarily add up to 100%. As the total number of empty cells in the table were of a high number a chi square analysis was not carried out, this was due to the results would not being meaningful in cases where there are a large number of empty cells and the expected cell counts are less than 5.

For detection failures the main contributory factors identified was 'being in a hurry' (34.5%) and 'inexperience' (29.1%) of the PTW rider. The age group 0-18 (29.1%) made more detection failures than any other rider group according to the group size. A prognosis failure where the PTW rider was not expecting the other road user to make a manoeuvre was quite evenly spread for all factors. In terms of accident configuration right turn against (19.5%) conflicts had the highest number of occurrences in the data. Decision failures and diagnosis failures both had high values for the PTW rider being in a hurry and speeding. Diagnosis failures were particularly high for riders between the ages of 19-25 (37.9%) and also for PTWs with engine capacities above 250 cc (72.4%). Both decision failure and diagnosis failure types had a higher level of injury compared to the other accident groups. Decision failures (22.2%) and prognosis (26.6%) failures were the most prominent groups for main contributing PTW riders in terms of right turn against accident situations. The other road user did not make an emergency manoeuvre in 83% of the coded cases.

Table 57: PTW rider failure types with risk factors and other important factors

Factor	Detection N=55	Diagnosis N=29	Prognosis N=207	Decision N=36	Execution N=5	Overall N=7	N 339
Contributory factor							
Speed	14.5	27.6	5.3	55.6	20.0	42.9	15.0
Alcohol	0.0	0.0	0.0	0.0	0.0	42.9	0.9
Distraction	12.7	0.0	0.0	8.3	0.0	0.0	2.9
In a hurry	34.5	55.2	5.3	55.6	0.0	85.7	21.2
Inexperience	29.1	10.3	2.9	5.6	40.0	100.0	10.3
Age range							
0-18	29.1	13.8	9.7	19.4	20.0	28.6	14.7
19-25	9.1	37.9	15.5	22.2	20.0	28.6	17.4
26-45	29.1	34.5	32.9	33.3	40.0	42.9	32.7

Factor	Detection N=55	Diagnosis N=29	Prognosis N=207	Decision N=36	Execution N=5	Overall N=7	N 339
46-65	3.6	6.9	10.1	2.8	0.0	14.3	10.9
66+	1.8	0.0	3.4	2.8	0.0	0.0	2.7
Missing	5.5	6.9	28.5	22.2	20.0	0.0	21.5
Engine size							
⩽ 50cc	23.6	10.3	12.1	22.2	40.0	14.3	15.3
51> cc ⩽ 250	16.4	17.2	15.9	44.4	0.0	42.9	16.5
Cc > 250	50.9	72.4	48.8	52.8	60.0	42.9	51.6
Missing	5.5	0.0	23.2	16.7	0.0	14.3	17.1
Day/Night							
Day	83.6	79.3	67.1	88.9	60.0	71.4	73.5
Night	10.9	17.2	22.2	8.3	40.0	28.6	18.9
Missing	1.8	3.4	10.6	5.6	0.0	0.0	7.7
Other road user emergency manoeuvre							
Yes	9.4	20.7	16.5	22.9	50.0	28.6	17.3
No	90.6	79.3	83.5	77.1	50.0	71.4	82.7
Injury severity							
Fatal	9.1	10.3	2.9	13.9	0.0	28.6	5.3
Serious	23.6	34.5	23.2	36.1	40.0	28.6	26.0
Slight	25.5	48.3	56.5	38.9	40.0	57.1	53.7
Non-injury	5.5	3.4	7.7	8.3	20.0	0.0	7.7
Level of involvement							
Primary	85.5	89.7	6.3	61.1	40.0	100.0	36.3
Secondary	3.6	0.0	10.1	25.0	20.0	0.0	9.7
Not contributory	7.3	10.3	83.6	2.8	40.0	0.0	54.0
Accident type							
Other vehicle right turn against	5.5	0.0	26.6	22.2	0.0	0.0	19.5
Rear-end	41.8	3.4	4.8	5.6	20.0	0.0	10.9
Both vehicles turning right	7.3	0.0	9.2	2.8	0.0	0.0	7.1
Merging roads	1.8	0.0	8.2	2.8	0.0	0.0	5.6
Drifting into opposite lane	5.5	24.1	1.0	2.8	20.0	57.1	5.3

Table 58 shows the different PTW failure types and manoeuvres for each specific accident failure type for multiple vehicle accidents that include PTWs from the OTS dataset. Only 301 of the 449 cases had manoeuvres coded, by the OTS analyst, and are included in the table below. The largest groups of failures were prognosis failures (n=163) and the largest accident type for these failures was turning accidents (55.8%). Overtaking accidents (22.0%) was the second largest manoeuvre types for this failure. The second largest failure group were detection failures (n=52) and the largest accident type was rear-end accidents (51.9%), then overtaking (21.1%) and turning accidents (15.4%). Overtaking was a particularly large group of accidents where diagnosis failures (44.4%) and decision failures (25.7%) occurred.

Table 58: PTW rider failure type and accident type

Accident Type	Detection N=52	Diagnosis N=27	Prognosis N=163	Decision N=35	Other N=24
Overtaking	21.1%	44.4%	22.0%	25.7%	8.3%
Loss of control	9.6%	18.5%	4.9%	8.6%	29.1%
Rear-end	51.9%	3.7%	8.4%	14.3%	4.1%
Turning	15.4%	7.4%	55.8%	22.9%	4.5%
Other	2.0%	26.0%	8.9%	28.5%	54.1%
Total	100%	100%	100%	100%	100%

Table 59 shows the different PTW rider failure types and manoeuvres for single PTW accidents. The largest group of failures were diagnosis failures (n=34), while the other failures were quite evenly split between the other 5 groups; detection failures (n=17), execution failures (n=17), overall failures (n=17), decision failures (n=16) and prognosis failures (n=9). The most common accident types were loss of control accidents either in straight ahead situations (n=47) or while turning (n=42). Diagnosis failures were particularly high for loss of control accidents, accounting for 27% of all single PTW accidents. Diagnosis failures (31.2%) together with execution failures (14.5%) and overall failures (13.6%) accounted for over half of the single PTW accidents. Other than loss of control accident types only overtaking (5.4%) and hitting obstruction accidents (5.4%) were observed to be the accident types that occurred for single PTW accidents.

Table 59: PTW single vehicle rider failure with accident type

Failure	Detection N=17	Diagnosis N=34	Prognosis N=9	Decision N=16	Execution N=17	Overall N=17	Total 109
Overtaking	11.8%	2.9%	0.0%	18.8%	0.0%	0.0%	6
LOC (straight)	29.4%	38.2%	55.6%	43.8%	47.1%	52.9%	47
LOC (turning)	29.4%	50.0%	22.2%	25.0%	47.0%	35.3%	42
Hitting obstruction	5.9%	2.9%	11.1%	6.3%	5.9%	5.9%	6
Other	23.5%	6.0%	11.1%	6.1%	0.0%	5.9%	8

Table 60 shows all incidents where the road user in conflict with the PTW rider was coded as making a detection failure. These types of accidents have been highlighted in the literature as ‘looked but did not see’ accidents. The main type of failure that the road user made was due to hurried information acquisition (35.3%), the second due to visibility constraints (30.6%), and the third due to focusing on another road side component (22.5%).

Table 60: Other road user PTW detection errors by failure subgroup type

Failure type	N	Percentage
Visibility constraint conditions	53	30.6
Information acquisition focused on another component	39	22.5
Hurried information acquisition	61	35.3
Interruption in information acquisition	8	4.6
Neglecting the need to search for information	12	6.9
Total	173	100

8.3.2 Cluster analysis

Cluster analysis factors

Table 61 shows all of the associated risk factors used in the cluster analysis. A total of 13 specific variables were selected to be entered into this analysis according to the most relevant risk factors that previous research had identified as present in PTW accidents. The variables were divided into four groups. The human factors selected were the main failure that the PTW rider was coded as making as well as the first contributory factor that was coded for the rider. The rider's level of involvement in the accident was also coded, determining what level of contribution the rider made to the accidents conflict situation. The main contributing factor that the rider made was coded with 10 specific groups. The interacting road users main failure (or single accident) was coded in order to identify the direct failure interactions, to clarify the type of multi-vehicle accidents or whether the PTW was involved in a single vehicle accident. The age group and gender of the rider were also included in the analysis

With regards to the PTW, the size of the engine was used to determine the types of PTWs that faced different obstacles. The PTW types were classified similarly to Montella, Aria, D'Ambrosio, & Mauriello (2012) where the first type referred to commonly as an L1 lightweight PTWs (mopeds and scooters) with a cylinder capacity less than or equal to 50 cm3 (category 1), and the second and third categories were made up of L3 categories scooters and light weight motorcycles, with a cylinder capacity greater than 50 cm3 and less than or equal to 250 cm3 (category 2) and the last category included heavy scooters and motorcycles with a cylinder capacity greater than 250 cm3 (category 3).

In terms of the environment and infrastructure separate factors that described the road area, speed limit and road type were entered into the analysis. The manoeuvre of the PTW rider was also included in the analysis as either going ahead on a straight road, entering an intersection, overtaking another vehicle and the remaining accident manoeuvres were put into a group called "other" as their numbers were not significant enough to be entered into the cluster analysis. A detailed list of all of the values counts and percentages can be found in Appendix B (pp. 353).

Table 61: Variables used in the PTW accident cluster analysis

Variable	Aspect	Level	Value
Speed limit	Environmental	Accident	$\geqslant$ 30 mph; 40-50 mph; 60-70 mph
Road area	Environmental	Accident	Urban; Rural
Light conditions	Environmental	Accident	Day; Night

PTW rider failure mechanism	Traffic accident	PTW rider	Detection; Diagnosis; Prognosis; Decision; Execution; Overall
Gender	Road User	PTW rider	Male; Female
Age group	Road User	PTW rider	0-18; 19-25; 26-45; 46-65; 66+
Rider contributory factor	Accident	PTW rider	Physical/physiological; Risk taking; Inexperience; Distraction; Road condition; Traffic condition; Visibility impaired; Other environmental factors; Vehicular factor; No factor
Emergency manoeuvre	Accident	PTW rider	Yes; No
Level of involvement	Accident	PTW rider	Primary contributory; Secondary contributory; Not contributory
PTW type	Vehicle	PTW rider	$\leqslant$ 50cc; 51> cc $\leqslant$ 250; cc > 250
Road type	Environmental	Accident	A class; B class; Motorway; Minor
Other vehicle failure type	Vehicle	Road User 2	Detection; Prognosis; Decision; Single PTW crash; Other
Rider manoeuvre	Accident	PTW rider	Leaving lane; Rear-end; Changing lane; Overtaking; Right turn; Left turn; Intersection; Other

Goodness of Fit analysis

The results from the goodness of fit analysis comparing the Akaike information criterion (AIC) and Bayesian information criterion (BIC) from 2 to 15 clusters can be seen in figure 24. The BIC (10919.30) produced a two cluster fit as the best fit for analysis purposes while the AIC (10657.58) produced a seven cluster fit as the best fit for analysis purposes. The AIC analysis was selected for the cluster analysis as the case size and dataset structure was a better fit for this analysis (figure 24).

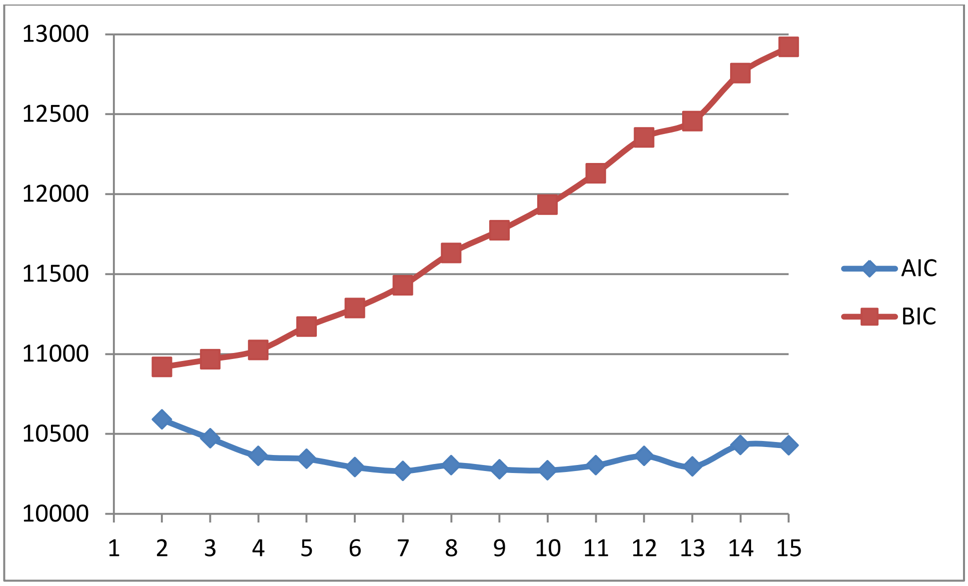


Figure 24: AIC and BIC values for the PTW accident cluster analysis

Cluster analysis descriptions

The latent class cluster analysis focused on the documented 428 PTW accident files. Seven distinctive (separated) accident classes were highlighted resulting in a 7 cluster solution. The grouping was made according to the cluster sizes, the largest cluster being the first group, the second largest being the second group and so forth (figure 25). A detailed table that includes all of the cluster results can be found in Appendix B (pp. 355), in this table each overly represented significant factor is presented in bold.

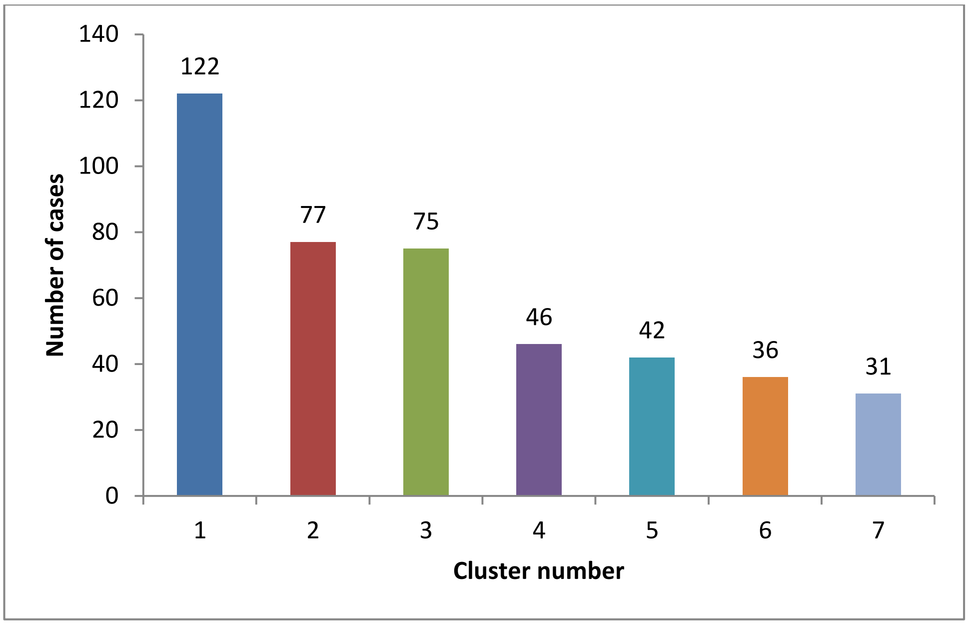


Figure 25: PTW cluster sizes

A simple explanation of each of the clusters was carried out in order to support the discussion, these explanations can be found within the results section. For each of the clusters a table that included all of the factors that were significant and over-represented according to the chi square analysis, as well as the degrees of freedom (df) values and number of cases where these factors were present was created. The degrees of freedom values and values of significance were not reported by each individual factor within the clusters to prevent repetition.

Cluster analysis results

Cluster 1 (n=122)

"Intersection accident due to other road user not detecting PTW"

Table 62 highlights the results for cluster 1, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 62: Powered two wheeler cluster 1 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Rider gender	Female	17.0	1	21	8.4	0.01
PTW rider failure mechanism	Prognosis	96.6	5	118	161.2	0.001
Rider contributory factor	Traffic condition	40.5	9	49	120.7	0.001
	No Factor	43.0	9	52		
Level of involvement	Not contributory	92.9	2	113	187.4	0.001
Other vehicle failure type	Detection	76.4	4	93	139.2	0.001
	Decision	14.0	4	17		
	Other	9.7	4	12		
PTW engine size	50cc	24.5	2	30	11.1	0.01
Road area	Urban	87.8	1	107	46.8	0.001
Speed limit	30 mph and under	76.0	2	93	79.9	0.001
Road type	B class	22.9	3	28	25.0	0.001
	Minor	39.8	3	48		
Accident situation	Right turn	53.6	7	65	114.6	0.001
	Left turn	6.4	7	8		
	Intersection	12.9	7	16		

Human Factors

PTW: In terms of demographic variables female riders (17.0%) were significantly over-represented for this cluster. The riders main failures were prognosis failures (96.6%) and the contributing factors for the rider were either the condition of the traffic environment (40.5%) or no contributory factor (43.0%) being coded. The fact that the rider made a prognosis failure and was not contributing (92.9%) to the accident highlights that most of the failures were related to the rider not expecting the other road user to make a manoeuvre. The failure that the other road user made was a detection failure (76.4%) or a decision failure (14.0%).

Vehicular Factors

The engine size of the PTWs were 50cc and below (24.5%).

Environmental/Infrastructural Factors

These accidents occurred in an urban area (87.8%) with a speed limit that was 30 mph or under (76.0%) in a B class (22.9%) or minor road (39.8%). The accident situation that the rider was going against was either a vehicle turning against right turn (53.6%), vehicle making a left turn (6.4%) or an intersection (12.9%).

Cluster 2 (n=77)

"Single vehicle PTW accident due to risk taking and incorrect diagnosis of the roadway"

Table 63 highlights the results for cluster 2, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 63: Powered two wheeler cluster 2 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Rider age group	19-25	34.8	4	27	26.2	0.001
PTW rider failure mechanism	Diagnosis	50.8	5	39	204.4	0.001
	Decision	25.1	5	19		
	Overall	21.2	5	16		
Rider contributory factor	Physical/physiological	33.5	9	26	110.1	0.001
	Risk taking	57.8	9	44		
Level of involvement	Primary contributory	100.0	2	77	94.5	0.001
Other vehicle failure type	Single vehicle	73.4	4	57	161.2	0.001
Engine size	250+cc	75.8	2	58	9.9	0.01
Road area	Rural	57.9	1	45	16.7	0.001
Speed limit	60-70 mph	46.2	2	36	24.9	0.001
Road type	Minor	44.3	3	34	11.1	0.05
Accident situation	Leaving lane	83.7	7	64	183.1	0.001

Human Factors

PTW: The PTW riders were male riders (90.9%) for this cluster despite gender not having a significant chi square value and the significant age range was 19-25 (34.8%).

The riders main failures were diagnosis failures (50.8%), decision (25.1%) or overall failures (21.2%). The contributing factors for the rider were either the physical/physiological condition (33.5%) or risk taking (57.8%). The rider was the primarily contributing road user (100.0%) to the accident. The accident was a single vehicle accident (73.4%).

Vehicular Factors

The engine size of the PTWs was above 250 cc (75.8%).

Environmental/Infrastructural Factors

These accidents occurred in a rural area (57.9%) with a speed limit that was above 60 mph (46.2%) in a minor road (44.3%). The rider was leaving their lane (83.7%).

Cluster 3 (n=75)

"Detection conflict situation with road user due to lane changing or turning"

Table 64 highlights the results for cluster 3, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 64: Powered two wheeler cluster 3 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Rider age group	66+	8.2	4	6	15.5	0.01
PTW rider failure mechanism	Prognosis	96.1	5	72	84.2	0.001
Rider contributory factor	Traffic condition	34.0	9	25	55.7	0.001
	No Factor	46.4	9	35		

Level of involvement	Not contributory	83.1	2	62	71.6	0.001
Rider emergency manoeuvre	Yes	46.5	2	35	4.1	0.05
Other vehicle failure type	Detection	79.5	4	60	73.5	0.001
Engine size	250+	78.2	2	59	9.7	0.01
Road area	Rural	74.0	1	55	51.6	0.001
Speed limit	40-50 mph	50.2	2	38	76.5	0.001
	60-70 mph	49.8	2	37		
Road type	A class	67.9	3	51	36.0	0.001
	Motorway	13.5	3	10		
Accident situation	Changing lane	31.4	7	28	72.8	0.001

Human Factors

PTW: Male riders were not significant for this cluster but were involved in a large proportion of the cases (92.6%). The rider's age range 66 years or older (8.2%) was over-represented.

In this cluster prognosis failures (96.1%) were identified as being significant and the contributing factors for the rider were either the condition of the traffic environment (34.0%), or no contributory factor was coded (46.4%). The rider was not contributing (83.1%) to the accident occurring and made an emergency manoeuvre (46.5%). The failure that the other road user made was a detection failure (79.5%).

Vehicular Factors

The engine size of the PTWs was above 250cc (78.2%).

Environmental/Infrastructural Factors

These accidents occurred in a rural area (74.0%) with a speed limit that was 40 mph to 50 mph (50.2%) or over 60 mph (49.8%) in an A class road (67.9%) or motorway (13.5%). The accident occurred while changing lanes (31.4%).

Cluster 4 (n=46)

“Rider detection issues in high speed situation”

Table 65 highlights the results for cluster 4, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 65: Powered two wheeler cluster 4 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Rider age group	46-65	28.5	4	13	18.2	0.01
PTW rider failure mechanism	Detection	74.8	5	34	98.3	0.001
	Diagnosis	25.3	5	11		
Rider contributory factor	Distraction	12.8	9	6	22.3	0.01
Level of involvement	Primary contributory	100.0	2	46	28.3	0.001
Other vehicle failure type	Prognosis	80.4	4	36	65.4	0.001
Engine size	250+ cc	84.2	2	38	12.7	0.01
Accident situation	Rear-end	41.8	7	19	84.6	0.001
	Overtaking	29.0	7	13		

Human Factors

PTW: Male riders (91.1%) were not significant for this cluster despite being involved in a high percentage of these cases. The rider’s age range was between 46-65 (28.5%).

The riders main failures were detection failures (74.8%) or diagnosis failures (25.3%), and the contributing factors for the rider were the rider’s distraction (12.8%). The rider was the primarily contributing road user (100%). The failure that the other road user made was a prognosis failure (80.4%).

Vehicular Factors

The engine size of the PTWs was above 250 cc (84.2%).

Environmental/Infrastructural Factors

The accident type that was significant was a rear-end accident (41.8%) or an overtaking accident (29.0%).

Cluster 5 (n=41)

"Young rider detection issues in low speed situations due to risk taking or inexperience"

Table 66 highlights the results for cluster 5, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 66: Powered two wheeler cluster 5 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Rider age group	0-18	88.0	4	37	168.8	0.001
PTW rider failure mechanism	Detection	53.8	5	23	70.1	0.001
	Overall	13.2	5	6		
Rider contributory factor	Physical/physiological	35.7	9	15	58.2	0.001
	Inexperience	16.7	9	7		
	Visibility impaired	9.3	9	4		
Level of involvement	Primary contributory	93.1	2	39	35.7	0.001
Other vehicle failure type	Prognosis	64.1	4	27	53.9	0.001
Engine size	50 cc	66.6	2	28	98.0	0.001
Road area	Urban	90.0	1	38	15.0	0.001
Speed limit	30 mph and under	84.0	2	35	29.8	0.001
Road type	Minor	58.8	3	25	20.9	0.001
Accident situation	Rear-end	22.0	7	9	14.8	0.05

Human Factors

PTW: Male riders were involved in a high proportion of these cases but were not over-represented for this cluster (90.5%), and the riders age range was between 0-18 (88.0%).

The riders main failures were detection failures (53.8%) and the contributing factors for the rider were either physical/psychological (35.7%), inexperience, (16.7%) or impaired visibility (9.3%). The rider was the primary contributing road user (93.1%). The failure that the other road user made was a prognosis failure (64.1%).

Vehicular Factors

The engine size of the PTWs was below 50 cc (64.1%).

Environmental/Infrastructural Factors

These accidents occurred in an urban area (90.0%) with a speed limit that was under 30 mph (84.0%) in a minor road (58.8%). The accident type was a rear-end accident (22.0%).

Cluster 6 (n=36)

“Young PTW rider with small engine size accident in a low speed setting”

Table 67 highlights the results for cluster 6, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 67: Powered two wheeler cluster 6 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
PTW rider failure mechanism	Decision	62.8	4	23	102.2	0.001
Rider contributory factor	Risk taking	49.9	9	18	31.6	0.001
Level of involvement	Secondary contributory	34.1	2	12	51.2	0.001
Other vehicle failure type	Decision	62.8	4	23	21.2	0.001
Road area	Urban	88.5	1	32	11.4	0.001
Speed limit	40-50 mph	48.5	2	17	2.5	0.01
Road type	A class	67.6	3	24	6.4	NS
Accident situation	Right turn	49.2	7	18	25.6	0.001

Human Factors

PTW: Male riders (95.5%) were not significant for this cluster. The rider's failure identified by the cluster analysis was decision failures (62.8%) and the contributing factors for the rider was risk taking (49.9%). The rider as the primary contributing (62.7%) road user had a high proportion but did not have a significant residual value. The failure that the other road user made was a decision failure (62.8%).

Vehicular Factors

No engine size value was significant in this cluster.

Environmental/Infrastructural Factors

These accidents occurred in an urban area (88.5%) with a speed limit that was either 40 or 50 mph (48.5%) in an A class road (67.6%). The accident situation was a right turn conflict (49.2%).

Cluster 7 (n=31)

"Single motorcycle leaving lane due to road condition or vehicular failure"

Table 68 highlights the results for cluster 7, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 68: Powered two wheeler cluster 7 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Rider age group	26-45	76.3	4	24	14.8	0.01
PTW rider failure mechanism	Execution	56.6	5	18	209.7	0.001
Rider contributory factor	Road condition	34.5	9	11	260.5	0.001
	Other environmental factors	18.8	9	6		
	Vehicular factor	15.7	9	5		
Level of involvement	Primary contributory	81.0	2	25	13.5	0.01
Other vehicle failure type	Single vehicle	74.2	4	23	61.7	0.001
Road area	Rural	64.8	1	20	10.6	0.01

Speed limit	60-70 mph	61.2	2	19	26.3	0.001
Road type	Motorway	31.8	3	10	50.3	0.001
Accident situation	Leaving lane	77.2	7	24	53.7	0.001

Human Factors

PTW: Male riders were not significant for this cluster though were in a high proportion of the accidents (96.6%). The rider's age range was between the ages of 26 to 45 (76.3%). The riders main failures were execution failures (56.6%) and the contributing factors for the rider were road conditions (34.5%), other environmental factors (18.8%) or vehicular factors (15.7%). The rider was the primarily contributing road user (81.0%). The accident was a single vehicle accident (74.2%).

Vehicular Factors

The engine size of the PTWs was not significant for this cluster.

Environmental/Infrastructural Factors

These accidents occurred in a rural area (64.8%) on a motorway (31.8%) with a 60 mph or over speed limit (61.2%). The accident type was a leaving lane accident (77.2%).

8.4 Discussion

The PTW study was undertaken with two general aims, the first was to understand the different types of accident scenarios that occur for PTW riders with other road users. The second was to determine the nature of these scenarios and what countermeasures can be identified to either reduce or altogether stop these types of accidents.

This analysis process was carried out in two parts, firstly carrying out a descriptive analysis in order to provide a detailed understanding of the accidents and what type of road users the PTW riders mostly interacted with, and secondly a cluster analysis in order to use a statistical approach to form scenarios based on inferential statistics.

8.4.1 Sampling

The sample in this study consisted of 449 cases collected by the OTS in-depth accident study between the years 2000-2010. McCarthy, Walter, Hutchins, and Tong (2008) carried out an analysis comparing the OTS PTW accidents against the relevant Great Britain national data cases, using the 302 PTW accidents that were collected up to that point (90 single PTW accidents, 212 multi-vehicle accidents). This comparison was carried out using a chi square analysis and found that the OTS data was not significantly different than national data with regards to rider age, engine size and types of area.

A comparison with regards to accident injury severity found that there was a major difference between the data, and that the OTS data had more accident cases that included severe injuries. The reason for this may be that the nature of in-depth accident studies make there accident collection procedure skew towards injury accidents, as the accidents are notified to police immediately and some non-injury or less significant injury accidents are not reported (Clarke et al., 2004).

8.4.2 Descriptive analysis

In terms of the accident configurations most of the accidents found in this dataset were either single vehicle accidents (24.3%) or multiple vehicle accidents that occurred with cars as the vehicle that the PTW was in conflict with (60.0%). This is in accordance with UK national data as reported by Elliot et al. (2003), but single PTW accidents were slightly more (18%) and car to PTW accidents (68%) less than figures reported in the UK by the DfT in 2005 (DfT, 2005).

In terms of demographic variables males were coded as the PTW rider in nearly 90% of these accidents which is similar to other studies on PTWs (Bjørnskau et al., 2012; MAIDS, 2009).

The single PTW accident causes were a result of the PTW either losing control on a straight road (52.9%) or losing control while turning (35.3%), this is in contrast to figures from Clarke et al. (2007) and Hurt et al. (1981) where both studies had most of the loss of control studies occurring on bends rather than on straight roads, though the figures from Hurt et al. (1981) are for US national data and the sample is not similar to this study, due to the years the data was collected and sample characteristics. While for multiple vehicle accidents the main accident types were right turn against (19.5%) and rear-end accidents (10.9%) which is similar to figures from Clarke et al. (2007). The data from the Clarke et al. (2007) study was obtained from the midland police forces between the years 1997 – 2002, so there may be some overlap between a small number of cases that were obtained in the Nottinghamshire region of the OTS study and similarities in the sample characteristics.

In terms of contributing to an accident's occurrence other road users (54%) were more likely than PTW riders (36%) to cause the conflict situation to occur. This is different compared to the data from Clarke et al. (2007), which had PTW users as the main contributing road user in 51% of the cases. The possible reasons for this could be that though one of the main sampling areas was similar to the OTS studies, the data was collected in different years and accident configurations may have changed during the different periods. Clarke et al. (2007) did not include non-injury cases and this may have also caused the difference. Another possible reason could be that the coding schema or the interpretation by the researchers used by the UK national data police force and OTS studies was different within the two studies.

Most of the accidents in this sample occurred in situations where the PTW rider was going straight ahead and was not in direct conflict with the other road user. The other road user did not make an emergency reaction on more than 80% of the two road user accidents.

8.4.3 Cluster analysis interpretation

The analysis classified the accidents into seven different clusters, from these clusters it was possible to differentiate the four clusters where the PTW rider was the main contributing road user to the accident occurring, two clusters where they were not a contributing user and one cluster where both road users contributed to the accident occurring. A brief description for each cluster is made below. A detailed listing of all relevant factors present in the

PTW cluster model as well as some explanatory descriptive variables for the clusters is present in table 69.

Cluster 1

The first cluster identified an accident situation where the road user was making a right turn at an intersection in a low speed setting and was not able to detect the PTW rider. The PTW rider did not expect to be not detected and so did not have sufficient time to react to this occurrence, and did not make an emergency manoeuvre. The main contributing road user to these accidents was the other road user. This accident can be termed a 'looked but did not see' right of way accident, which can be classified as right of way violation accidents at T junctions identified by Clarke et al. (2007). This group of car to PTW accidents was the accident that most commonly occurred.

Cluster 2

Cluster 2 identified a single road user accident where the PTW rider misdiagnosed the roadside and made a manoeuvre that led to them leaving the lane. This cluster was related to risk taking and the rider did not make an emergency manoeuvre. These accidents were similar to the accidents classified by Clarke et al. (2007) as loss of control accidents with the accident setting described being the same as this cluster.

Cluster 3

The third cluster identified a situation where the road user was changing lanes or turning right in a medium to high speed setting and did not detect the PTW rider. The other road user was the main contributing road user to this cluster. The PTW rider did not expect the other road user to perform their behaviour but was able to make an emergency manoeuvre on nearly half of the cases. This cluster could also be determined as 'looked but did not see' accidents though different to cluster 1, in that the there was a higher speed setting and the size of the PTW was larger.

Cluster 4

Cluster 4 identified an accident where a rider on a large engine PTW made a detection or diagnosis error. The cases for this cluster were distributed between either rear-end or overtaking accidents, where the manoeuvre that the PTW rider made led to the accident occurring.

Cluster 5

Cluster 5 identified accidents where younger riders were on PTWs that were 50ccs or below, and made a detection error due to risk taking or inexperience. The accident configuration was the rider either not seeing the vehicle in front and striking the other vehicle in the rear or performing an unsuccessful overtaking manoeuvre. These accidents were similar to the accidents classified by Clarke et al. (2007) as PTW manoeuvring accidents and the group of young moped riders were far more likely to cause rear-end shunts compared to other PTW riders.

Cluster 6

This cluster identified accidents where both road users were contributing to the accidents occurrence. The main types of accidents that were highlighted were right turn accidents, where the rider made a decision error due to risk taking and the other road user made a detection or decision failure. This cluster occurred in urban areas and younger PTW riders were over-represented for this grouping.

Cluster 7

This cluster identified single vehicle PTW accidents where the PTW rider either made an overall failure or an execution failure. The road condition was deemed to be problematic in a quarter of the cases and there was a PTW vehicular failure in another 20% of these cases. This was the only cluster that had the PTW as the primary contributing road user together with traffic condition as a contributory factor in a multiple road user accident.

Cluster analysis general discussion

The cluster analysis mainly distinguished between accidents where the PTW was the main contributing road user and accidents where the other road user was the main contributing road user.

For the clusters where the road user was the main contributing road participant cluster 1 and cluster 3 highlighted detection issues in two different settings. These clusters are ‘looked but did not see accidents’ where the other road user made a detection error. This conflict situation between PTWs and vehicle types, particularly car to PTW, has been illustrated by previous research (Brown, 2002; Clabaux et al., 2012; Clarke et al., 2004; Clarke et al., 2007; Crundall et al., 2012; Hurt et al., 1981) and is of particular importance for road safety measures to broach as a large number of cases in research all over the world all demonstrate this accident type.

The difference between the clusters is the first cluster being in a low speed setting in an urban area where there was an issue with the traffic environment with a right turn against or intersection situation. The third cluster contained cases that were in a higher speed setting and rural environment which is different from typical ‘looked but did not see’’ accidents.

Koustanai et al. (2008) identified that road users failed to see a two-wheeled vehicle mainly because it has atypical properties in comparison with the rest of traffic. The reason that PTW riders are not detected by other road users can be many, similar results have been discussed by previous studies that identified ‘looked but did not see’ PTW accidents, but the in-depth accident analysis provided in this chapter that further separated these failures can help in providing a discussion point of the different types of detection errors. The descriptive analysis in table 60 highlighted hurried information acquisition as the main issue in these types of accidents. The other main types of failures were due to visibility constraints in the environment and focusing on another component within the traffic environment.

Table 69: PTW cluster analysis model and variables

Cluster/ N of cases/% of cases	Descriptive information		Cluster model						
	Accident type	Casualty level	PTW Main failures	Contributory factor 1	Other road user failure	Gender/ Age	Manoeuvre	Level of involvement	Accident setting
1 122 28.5%	Right turn accident onto PTW	Fatal (1) Serious (28) Slight (68) Non-injury (9)	Prognosis	Traffic condition	Detection	Female/ 0-25	Right turn Left turn Intersection Other	Not contributory	B road/ Minor road 30 mph Urban area Night-time
2 77 18.0%	PTW loss of control	Fatal (15) Serious (31) Slight (27) Non-injury (3)	Diagnosis Decision	Risk taking Physical/Psychological	Single vehicle accident	Male/ 19-45	Leaving lane	Primary contributory	Minor road 60-70 mph Rural area Daytime
3 75 17.5%	Loss of control on a straight road or a bend	Fatal (3) Serious (17) Slight (46) Non-injury (6)	Prognosis	Traffic condition	Detection	Male/ 26-45	Changing lane Intersection	Not contributory Secondary contributing	A road/ Motorway 40-50 mph /60-70 mph Rural area Daytime
4 45 10.5%	Rear-end/ Overtaking	Fatal (2) Serious (16) Slight (23) Non-injury (3)	Detection Diagnosis	Distraction	Prognosis	Male/ 26-65	Rear-end Overtaking	Primary contributing	A road Daytime
5 42 9.8%	Right turn/Rear-end	Fatal (1) Serious (7) Slight (30) Non-injury (3)	Detection Overall	Physical/Physiological Inexperience Visibility impaired	Prognosis	Male/ 0-18	Rear-end Intersection	Primary Contributing	Minor rad 30 mph Urban area
6 36 8.4%	Right turn	Fatal (5) Serious (14) Slight (15) Non-injury (1)	Decision	Alcohol Breaking the law	Detection Decision	Male	Right turn	Primary contributing Secondary contributing	A road 40-50 mph Urban area Daytime
7 31 7.2%	PTW leaving lane	Fatal (2) Serious (10) Slight (15) Non-injury (4)	Execution	Road condition Other environmental factors	Single vehicle Other	Male/ 26-45	Leaving lane	Primary contributing	Motorway 60-70 mph Rural area

Single PTW accidents were prominent in clusters 2 and 7. The first cluster had driver risk taking as a large contributory factor to the accidents occurrence, identifying that riders misdiagnosed the road situation in a high speed setting. In both of these clusters the road user age group that was significant was either 19-26 year olds or 26-45 year olds. Both of these clusters had risk taking as the main contributing factor though the accident settings were different in that the former was a high speed rural setting while the latter occurred on low speed urban roads and occurred at a conflict situation rather than a loss of control.

The second clusters difference compared to the first cluster was with regards to the engine size of the PTWs and being in a higher speed setting in a rural area with lane changing by the vehicles. Clarke et al. (2004) reported from a questionnaire study that 58% of motorcyclists admitted to always or frequently breaking the speed limit, and both of these clusters are related to younger PTW riders riding at speeds that are not suitable for the situation that they are in.

In cluster 7 the segmentation included environmental visibilities and the traffic environment as contributory factors and both road users contributing to these accident types.

8.4.4 Countermeasure indications

When considering traffic safety countermeasures it is possible to consider different stages of the accident where countermeasures can be attributed, these can be the pre-crash, crash or post-crash. As this study focused on how the accident occurred and not the injury outcomes of each individual accident the countermeasure discussion will focus on the pre-crash phase.

When discussing possible countermeasure developments it is necessary to break the different PTW accidents into three groupings based on the cluster analysis:

1. Single PTW accidents
2. Accident were the other road user was the primary contributing vehicle
3. Accidents were the PTW was the primary contributing vehicle

In the single PTW accidents, where the PTW was the primary contributing vehicle and made risk taking behaviours, namely cluster 2, a number of countermeasures can be considered based on education measures. Clarke et al. (2004) identified appropriate countermeasures for cases where motorcyclists initiated the accident and had the possibility of making a countermeasure. These countermeasures were identified as slower speed on bends, appropriate speed for conditions and not overtaking near a junction or exit.

Extra training or awareness programs for riders in traffic conditions that are wet or snowing may benefit riders in the long run. Anti-lock Braking Systems (ABS) specifically developed for PTWs is becoming the norm but older bikes do lock up when undertaking manoeuvres. ABS has become compulsory for PTWs above 125cc since 2016 in the EU and this is a good step forward to take, particularly for PTW performance on difficult and wet/icy roads. Also awareness with regards to risk taking would be of particular benefit to motorcyclists and this could be possible through either programs.

It also has to be considered that PTW active safety systems would be beneficial particularly for these single vehicle accidents, as they would either improve the riders ability to respond to the situation earlier by providing notification of a possible conflict or provide the opportunity for their emergency reaction whether it be breaking, swerving or both combined to be more successful by creating an optimum performance for each behaviour for the rider.

For the accidents where the other road user was the main contributing road user ’looked but did not see’ failures were prominent for clusters 1 and 3. From the data provided in these cases it can be seen that the road user is unaware of the PTW rider and so does not make any mitigating manoeuvre for the accident while the PTW rider also does not make a manoeuvre. Ideally in these situations the other road user could be made aware of the PTW rider. How to do this is an issue of concern as there are possible measures to introduce though the measures effectiveness are debateable.

One of the main issues is the visual search patterns of other road users in relation to riders. Crundall et al. (2012) demonstrated that car road users that have PTW licenses are more likely to see PTWs than car road users who do not ride PTWs, though the time spent for initial perception of the PTW was the same the appraisal of the rider was made far more quickly by the driver with riding experience. So, rather than scanning behaviours for the road user in order to identify PTW riders, effective scanning of the PTW rider in order to speed up the appraisal process of these road users is necessary. Ideally driver training would be carried out in this regard to help road users. The question of how to effectively provide this type of training is a difficult one to answer, as LBDNS car to PTW right of way accident have been studied and discussed throughout the last 30 years and are still occurring frequently.

The three main types of detection failures that were identified have different countermeasure possibilities. All of the failure types would benefit from technological developments helping the other road user to identify the rider. A number of systems are starting to be introduced into the market place and could provide benefits in this regards.

Scholliers, Bell, Morris, and Garcíad (2014) identified oncoming vehicle information systems for PTWs, based on vehicle to PTW communication, as a possible avenue of improvement which could provide other road users with information on PTWs and allow for the appraisal and diagnosis of conflict situations to run more smoothly.

Three of the clusters occurred on urban roads with lower speed limits, and thus as Clabuax et al. (2012) identified traffic calming measures could be used to specifically target these situations. One of the accident types involved an accident with small PTWs where young riders made errors due to their inexperience, their physiological/ psychological state or being distracted. The main countermeasure for this would be to either better train or educate these riders.

8.5 Summary

Chapter 8 described a statistical analysis of all of the PTW accident cases collected in the OTS study with regards to human failure and different accident factors. The study analysed 448 cases from the OTS dataset performing a descriptive analysis and 428 cases were entered into a latent class cluster analysis.

The results identified 7 scenarios with regard to PTW rider accidents, defining 5 scenarios where the PTW rider was the primary contributing road user and 2 scenarios where the other road user was the primary contributing road user. The main accident types where the road user was the contributing factor were 'looked but did not see' accidents, while the accidents where the PTW rider was the main contributing road user were accidents that occurred due to speeding or risk taking. The countermeasure indication for each scenario was discussed with regards to possible safety measures for the road users.

9 An Analysis of Pedestrian Accidents

9.1 Introduction

The literature review suggested that a better understanding of traffic accidents can be accomplished by using an interactive ergonomics model to understand how the multitude of factors combine to cause the main human failure. In traffic accidents certain factors have been identified as increasing the risk of an accident occurring. Pedestrian accidents are a particular group of accidents that are important for road safety performance improvement, as the road user is not protected by any vehicle.

Rather the immediate reaction is with the other object in the road and thus the injury outcomes of the accidents are usually greater than those for other accident types. Also, due to the injury outcomes of pedestrian accidents being more severe than other types of vehicle accidents, pedestrians are more dependent on other road users' behaviours and adherence to road safety rules.

The study reported here aimed to investigate the type of failure sequences that cause pedestrian accidents. This study analysed driver and pedestrian behaviour before pedestrian accidents and the sequence that leads up to these accidents.

9.2 Method

9.2.1 Design

Data were acquired from on the spot analysis of accident data by a group of accident researchers within a 30 minute time span of the accident occurring. Factors relating to the accident were obtained by grouping the accident variables into 4 specific groups relating to human, vehicular, infrastructural or environmental factors relating to the accident.

9.2.2 Sample

Accident data was selected from the On the Spot study collected between the years 2000 to 2010. As cluster analysis requires a large number of cases 10 years of data were included rather than the 4 years of data included in the full dataset analysis, as the number of pedestrian cases for this 4-year period was 111 cases which would have not allowed a large number of variables to be entered in the cluster analysis.

From this data 265 accidents involving pedestrians were analysed. All cases were included in the descriptive analysis. Of the cases that were analysed 17 did not have values for all of the variables that were included in the cluster analysis and so were omitted. In total 248 cases were included in the cluster analysis. Cyclists were also considered for analysis purposes but due to insufficient case numbers were not included as they did not fit the statistical requirements of cluster analysis.

In the sample 157 of the Pedestrians were male and 102 were female while 6 of the participant's gender was unknown. Of the 245 individuals whose age was coded in the sample the age for the pedestrians were on average 31.3 years old with a standard deviation of 24.2 (figure 26).

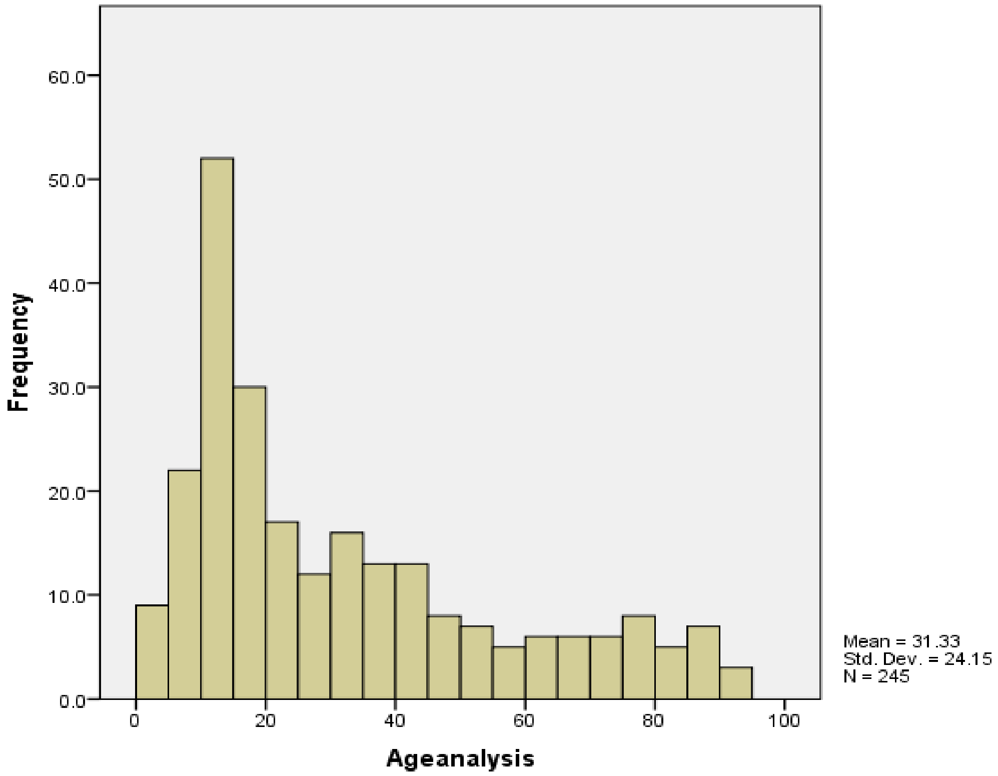


Figure 26: The age distribution of the pedestrian sample

9.2.3 Procedure

For this study the available accident data was analysed retrospectively by the author using the Human Functional Failure Causation methodology (Naing et al., 2007) and coded using the LAB accident type coding diagrams. All cases were then merged with the OTS data that was previously identified by OTS accident investigators. For each of the cases the pedestrian was identified as the first road user and the other vehicle was identified as the second road user.

9.2.4 Statistical analysis

This study incorporated a cluster analysis in order to group the accidents in different types of scenarios. This was done by separating the specific factors identified in table 74 and entering them into a cluster analysis. Each variable was grouped according to previous research and possible research questions, as the number of values that can be entered into analysis are limited. In the case that a number of values in the specific variable were small and could not be entered as a separate value they were combined into a value called other. This was done so that the small number of this value would not unduly influence the cluster analysis.

For this study a latent class cluster analysis based on the similarity between the different road users' involved in the accidents was used.

9.3 Results

9.3.1 Descriptive analysis

Below is a description of the type of impacting vehicles that interacted with pedestrians that were analysed within this chapter. Table 70 has the number of vehicle records for all of the accidents that are present within the analysed dataset. The first road user within the accident was always coded as the vehicle that directly interacted with the pedestrian. Within this dataset there were a total of 265 accidents coded. The majority of these cases were car to pedestrian accidents (80.4%).

Table 70: Vehicles included as the first interacting vehicle with pedestrians in the analysis

Interacting Vehicle	Count	Percent
Car	213	80.4
PTW	9	3.4
Bus	22	8.3
Heavy Goods	8	3.0
Light Goods	8	3.0
Other	5	1.9
Total	265	100.0

Table 71 illustrates the main failures made by the other road user and the pedestrian. For the other road user the main failure that was mostly coded is a prognosis failure (n=144), while the opposing main failures for the pedestrian road user were either making an overall failure (n=54), making a detection failure (n=41) or making a decision failure (n=39) due to undertaking risky behaviours. The other main type of failure for the other road user was detection failures where they could not identify the pedestrian (n=67). There was a high number of overall failures (n=75), related to impairment or fatigue, made by the pedestrian in this sample.

Table 71: Main failure types for pedestrians and other road users

Vehicle	Pedestrian						
Failure	Detection	Diagnosis	Prognosis	Decision	Execution	Overall	N
Detection	17	5	24	7	1	14	68
Diagnosis	6	0	5	2	0	3	16
Prognosis	41	3	4	39	3	54	144
Decision	1	1	11	3	1	2	19
Execution	0	0	4	0	0	1	5
Overall	2	2	8	0	0	1	13
Total	67	11	56	51	5	75	265

Table 72 contains a comparison of the injury outcome of the accident and the road type that was reported. Most of the accidents either occurred on an A class road (n=112) or a Minor road (n=99). Of the fatal or serious accidents 57 occurred on A roads, 42 on Minor roads and 17 on B roads.

Table 72: Main failure types for pedestrians and other road users

	Road Type				
Injury level	A road	B road	motorway	Minor road	Total
Fatal	13	4	0	5	22
Serious	44	13	1	37	95
Slight	52	23	1	52	138
Non-injury	3	0	1	5	9
Total	112	40	3	99	264

Table 73 contains a number of demographic, human factor and environmental/infrastructure factors against the failure that the pedestrian road user made. The pedestrian was the main contributory road user in the accident in 194 (73.2%) of the cases. The main contributory factors that were made by the pedestrian were being in a hurry (31.3%), playing (14.3%) and using alcohol (13.6%). This sample included a large number of pedestrians who were below the age of 18 (48.3%). Most of the accidents occurred during the day (69.8%) and in urban areas (70.2%). There were a large number of accidents that caused serious injuries (35.5%).

Table 73: Descriptive statistics of data against pedestrian failure type

Factor	Detection N=67	Diagnosis N=11	Prognosis N=57	Decision N=49	Execution N=5	Overall N=76	N 265
Pedestrian contributory factor							
In a hurry	70.1	36.4	1.8	55.1	40.0	56.6	31.3
Pedestrian playing	14.9	0.0	0.0	8.2	20.0	50.0	14.3
Alcohol	0.0	0.0	0.0	2.0	0.0	34.2	13.6
Traffic signs disobeyed	13.4	0.0	1.8	32.7	0.0	9.2	9.8
Eccentric behaviours	3.0	0.0	0.0	26.5	0.0	11.8	5.7
Visibility impaired	34.3	9.1	5.3	10.2	20.0	19.7	12.8
Distraction	4.5	18.2	0.0	10.2	80.0	21.1	6.8
Other road user contributory factor							
In a hurry	7.5	45.5	47.4	2.0	0.0	5.3	15.8
Risk taking	7.5	36.4	40.4	8.2	20.0	5.3	15.5
Visibility	32.8	9.1	17.5	30.6	20.0	36.8	29.1
Age range							
0-12	47.8	27.3	19.3	24.5	40.0	53.9	38.1
13-17	9.0	18.2	10.5	14.3	0.0	7.9	10.2
18-29	10.4	9.1	31.6	24.5	0.0	14.5	18.5
30-65	9.0	18.2	7.0	10.2	0.0	5.3	7.9
66+	11.9	0.0	21.1	14.3	0.0	7.9	12.5
Missing	11.9	27.3	10.5	12.2	60.0	10.5	12.8
Day/Night							
Day	77.6	72.7	73.7	67.3	80.0	60.5	69.8
Night	20.9	27.3	26.3	32.7	20.0	39.5	29.8
Injury severity							
Fatal	4.5	0.0	10.5	14.3	0.0	7.9	8.3

Serious	29.9	45.5	40.4	38.8	0.0	35.5	35.5
Slight	61.2	45.5	45.6	40.8	100.0	34.2	46.4
Non-injury	3.0	9.1	3.5	6.1	0.0	2.6	3.8
Road type							
A class	35.8	45.5	31.6	65.3	40.0	40.8	42.3
B class	14.9	9.1	21.1	10.2	40.0	13.2	15.1
Motorway	0.0	0.0	1.8	2.0	0.0	1.3	1.1
Minor	49.3	45.5	43.9	26.5	20.0	43.4	41.5
Level of involvement							
Primary	95.5	36.4	3.5	98.0	60.0	96.1	73.2
Secondary	3.0	0.0	0.0	0.0	40.0	3.9	2.6
Not contributory	1.5	63.6	96.5	2.0	0.0	0.0	24.2
Area type							
Urban	77.6	72.7	73.7	67.3	80.0	60.5	70.2
Rural	20.9	27.3	26.3	32.7	20.0	39.5	29.8

9.3.2 Cluster analysis

Cluster analysis factors

Table 74 illustrates all of the associated risk factors included in the analysis. A total of 13 specific variables were selected to be entered into this analysis according to the most relevant risk factors that are present in pedestrian accidents. The human factors selected were the main failure that the road user in conflict with the pedestrian was coded as making as well as their level of contribution to the accident. Whether the road user was contributing to the accident or not was also coded as the level of involvement of the road user. The pedestrian's main failure was also coded as well as the pedestrians contributory factor that contributed to the accident occurring. The age group and gender of both road users were also included in the analysis. In terms of the environment and infrastructure different factors that described the road type, speed limit and carriageway class were included. The driving behaviour of the road user as well as the conflict situation type was also included in this analysis. A detailed list of all of the values counts and percentages for these variables and values can be found in Appendix B (pp. 358).

Table 74: Variables used in the pedestrian accident cluster analysis

Variable	Aspect	Level	Value
Speed limit	Environmental	Accident	≥ 30 mph; <30 mph
Road area	Environmental	Accident	Urban; Rural
Light conditions	Environmental	Accident	Day; Night
Pedestrian failure mechanism	Traffic accident	Pedestrian	Detection; Diagnosis; Prognosis; Decision; Execution; Overall
Gender	Road user	Pedestrian	Male; Female
Age group	Road user	Pedestrian	0-12; 13-17; 18-29; 30-65; 66+
Gender	Road user	Road user	Male; Female
Age group	Road user	Road user	18-24; 25-34; 35-45; 46-65; 66+
Pedestrian contributory factor	Accident	Pedestrian	Alcohol/Impairment; Young age/Pedestrian playing; Psychological state; Risk taking; Driver behaviour; Visibility impaired; None
Pedestrian behaviour	Accident	Pedestrian	Crossing road; Crossing intersection; Crossing between cars; Vehicle crash; Other
Emergency manoeuvre	Accident	Road user	Yes; No
Level of involvement	Accident	Road user	Primary contributory; Secondary contributory; Not contributory
Road user situation	Accident	Road user	Going ahead; Traffic lights; Intersection; Overtaking; Pedestrian Crossing; Other
Other road user failure type	Vehicle	Road User	Detection; Diagnosis; Prognosis; Decision; Execution; Overall

Goodness of Fit analysis

The results from analysing goodness of fit measures in terms of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) and clusters ranging from 2 to 15 classes can be seen in figure 27. According to the BIC (7078.026) the selected solution was a three cluster model and according to the AIC (6581.895) the selection solution was a four cluster model. Due to the low number of cases and parsimony levels that were discussed in the methodology section a four cluster model was selected based on the AIC results for analysis purposes and used for the cluster analysis.

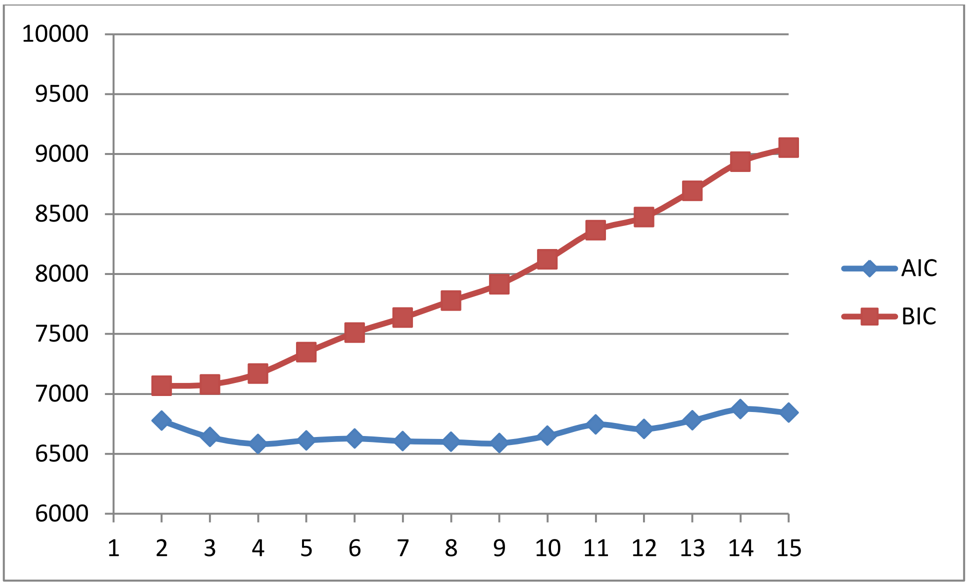


Figure 27: AIC and BIC values for the pedestrian cluster analysis

Cluster analysis

The latent class cluster analysis focused on the documented 248 other road user to pedestrian accident cases. Four distinctive (separated) accident classes were highlighted resulting in a 4-solution cluster, the clusters were put in order from largest to smallest. The number of cases for each cluster can be seen in figure 28. A detailed table of all of the cluster results can be found in Appendix B (pp. 360), in this table each overly represented significant factor is presented in bold.

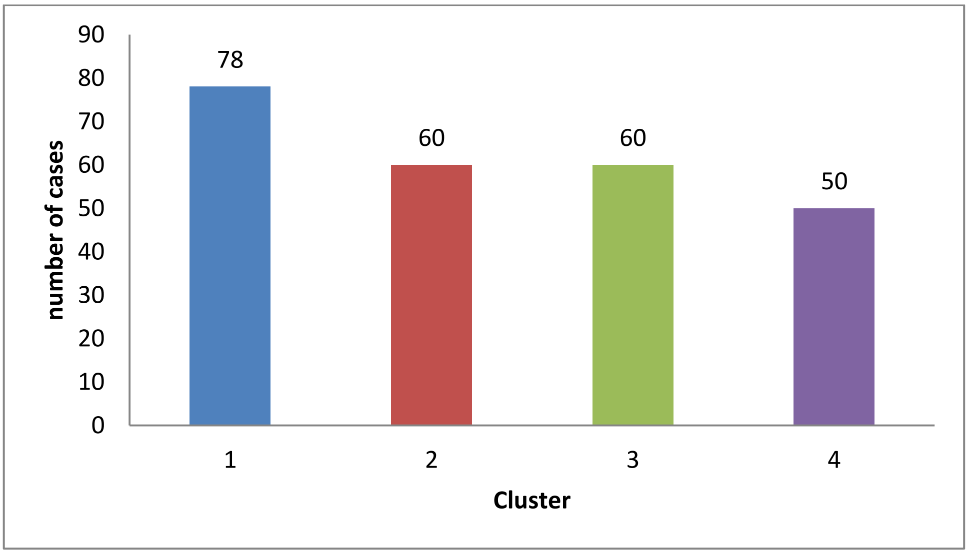


Figure 28: Pedestrian cluster sizes

A table describing the significant factors identified by the cluster analysis, their percentages in the cluster, number of cases they were equivalent to, degrees of freedom (df) values and values of significance were provided and an explanation of these factors was carried out in order to support the discussion, these explanations can be found within the results section.

Cluster analysis results

Cluster 1 (n=78)

‘’Pedestrians crossing an intersection while either under stress or under situations with limited visibility”

Table 75 highlights the results for cluster 1, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 75: Pedestrian cluster 1 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Pedestrian gender	Female	47.6	1	37	4.8	0.05
Pedestrian age	18-29	24.7	4	19	11.0	0.05
Pedestrian failure mechanism	Detection	66.2	5	52	150.6	0.001
Pedestrian contributory factor	Psychological state	59.2	6	46	134.0	0.001
	Visibility impaired	16.5	6	13		
Other road user age	26-45	37.7	4	29	19.3	0.001
Other road user failure type	Prognosis	73.1	5	57	25.6	0.001
Road user level of involvement	Not contributory	95.4	2	74	48.1	0.001
Light conditions	Day	83.1	1	65	9.8	0.01
Pedestrian behaviour	Crossing intersection	27.9	4	22	15.4	0.01

Human Factors

Vehicle: Road users aged between 26 and 45 (37.7%) were over-represented in this cluster. The failures that were significant for the other road user in this cluster were prognosis failures (73.1%). The other road user was not contributory to this accident occurring (95.4%).

Pedestrian: Female pedestrians were significant for this cluster (47.6%) and the pedestrian's age range was between 18-29 (24.7%). The pedestrian's main failures were detection failures (66.2%) and the contributing factors for the pedestrian were either their psychological state (59.2%) or their visibility being impaired (16.5%).

Environmental/Infrastructural Factors

These accidents occurred during the day (83.1%) and the pedestrian was crossing an intersection (27.9%).

Cluster 2 (n=60)

"Other road user with detection issues or in a hurry"

Table 76 highlights the results for cluster 2, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 76: Pedestrian cluster 2 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Pedestrian gender	Female	53.1	1	32	8.0	0.01
Pedestrian age	30-65	41.5	4	25	21.3	0.001
	66+	23.7	4	14		
Pedestrian failure mechanism	Diagnosis	10.0	5	6	198.6	0.001
	Prognosis	82.9	5	50		
Pedestrian contributory factor	Other driver	36.5	6	22	177.4	0.001
	None	50.2	6	30		
Other road user age	66+	23.2	4	14	7.7	NS
Other road user failure type	Detection	46.4	5	28	93.9	0.001
	Decision	16.6	5	10		
	Execution	6.6	5	4		

	Overall	16.6	5	10		
Road user level of involvement	Primary contributory	100.0	1	60	200.6	0.001
Other road user emergency manoeuvre	No	68.0	1	41	15.9	0.001
Accident situation	Pedestrian crossing	13.3	5	8	46.7	0.001
	Other	24.9	5	15		
Pedestrian behaviour	Vehicle crash	31.5	4	19	24.2	0.001
	Other	33.2	4	20		

Human Factors

Vehicle: Male road users (73.5%) were not significant in the chi square analysis despite their high numbers. Road users aged 66 and older (23.2%) were significant for this cluster. The failures that were significant for the other road user in this cluster were detection failures (46.4%), decision failures (16.6%), execution failures (6.6%) and overall failures (16.6%). The other road user was the primarily contributing road user to this accident occurring (100.0%). The other road user did not make an emergency reaction for this cluster grouping (68.0%).

Pedestrian: Female pedestrians were significant for this cluster (53.1%). The pedestrian's age range was between 30-65 (41.5%) and 66 years and older (23.7%).

The pedestrian's main failures were prognosis failures (82.9%) and diagnosis failures (10.0%). The contributing factors for the pedestrian were the other road user (36.5%) or no factor coded (50.2%).

Environmental/Infrastructural Factors

These accidents occurred where the other road user was either going ahead (59.3%), at a pedestrian crossing (13.3%) or coded as other accident (20.3%). The pedestrian conflict was either a result of a vehicle to vehicle accident (31.6%) or other accident type (33.2%)

Cluster 3 (n=60)

“Night time accidents with impaired or risk taking pedestrians while the pedestrian crosses the road”

Table 77 highlights the results for cluster 3, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 77: Pedestrian cluster 3 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Pedestrian gender	Male	82.2	1	49	10.4	0.01
Pedestrian age	19-29	27.9	4	17	45.7	0.001
	30-65	44.2	4	26		
	66+	22.3	4	13		
Pedestrian failure mechanism	Decision	39.4	5	24	56.4	0.001
	Overall	52.2	5	31		
Pedestrian contributory factor	Alcohol/Impairment	56.4	6	34	179.7	0.001
	Risk taking	33.0	6	20		
Other road user age	19-25	47.2	4	28	14.5	0.01
Other road user failure type	Prognosis	68.9	5	41	8.9	NS
Road user level of involvement	Secondary contributory	10.9	2	7	18.5	0.001
	Not contributory	81.2	2	49		
Light conditions	Night	57.7	1	35	28.1	0.001
Speed limit	Over 30 mph	19.6	1	12	64.8	0.05
Accident situation	Traffic lights	32.1	5	19	17.7	0.01
Pedestrian behaviour	Crossing road	47.8	4	29	24.2	0.001
	Crossing intersection	30.5	4	18		

Human Factors

Vehicle: Male road users were not significant for this cluster (79.9%). All road users in three age groups between nineteen and above were over-represented. The failures that were significant for the other road user in this cluster were prognosis failures (68.9%) which had a significant residual value despite the failure type grouping not having a significant chi square value. The other road user was not contributory to this accident occurring (81.2%).

Pedestrian: Male pedestrians were significant for this cluster (82.2%) and the pedestrians age range was between the ages of 19-29 (27.9%), 30-65 (44.2%) or 66 and older (22.3%). The pedestrian's main failures were decision failures (39.4%) or overall failures (52.2%) and the contributing factors for the pedestrian was either alcohol impairment (56.4%) or risk taking (33.0%).

Environmental/Infrastructural Factors

These accidents occurred during the night (57.7%) on a road with a speed limit that was over 30 mph (19.6%). The other road user was at traffic lights (32.1%). The pedestrian was either crossing the road (47.8%) or crossing an intersection (30.5%).

Cluster 4 (n=50)

"Young pedestrians crossing the road while playing or in a hurry"

Table 78 highlights the results for cluster 4, each significant and over-represented factor from the chi square analysis is described in detail in the section below.

Table 78: Pedestrian cluster 4 analysis results

Variable	Value	Percentage	df	N	χ^2	Sig.
Pedestrian age	0-12	70.3	4	35	86.6	0.001
	13-17	29.7	4	15		
Pedestrian failure mechanism	Overall	73.5	5	37	57.3	0.001
	Execution	4.0	5	2		
Pedestrian contributory factor	Young age/Pedestrian playing	77.1	6	39	122.2	0.001
Other road user failure type	Prognosis	69.2	5	35	7.4	NS
Road user level of involvement	Not contributory	86.3	2	43	15.7	0.001

Other road user emergency manoeuvre	Yes	75.5	1	38	11.4	0.001
Light conditions	Day	81.4	1	41	44.1	0.05
Accident situation	Going Ahead	82.8	5	41	32/7	0.001
	Overtaking	15.1	5	8		
Pedestrian behaviour	Crossing between cars	55.4	4	28	56.9	0.001

Human Factors

Vehicle: The other road users gender was not significant for this cluster, and the age groups variable had a significant chi square value but none of the groups were over-represented. The failures that were significant for the other road user in this cluster were prognosis failures (69.2%). The other road user was not contributory to this accident occurring (86.3%). The other road user made an emergency reaction for this cluster grouping (75.5%).

Pedestrian: Male pedestrians were not significantly over-represented in this cluster (70.2%) despite having a high percentage of the individuals coded. The pedestrians age range was between 0-12 (70.3%) or 13-17 (29.7%). The pedestrian's main failures were either overall failures (73.5%) or detection failures (4.0%), and the contributing factors for the pedestrian were young age/pedestrian playing (77.0%).

Environmental/Infrastructural Factors

These accidents occurred during the day (81.4%) with a speed limit that was at or under 30 mph (94.0%). The other road user was either going ahead (82.8%), or performing an overtaking manoeuvre (15.0%). The pedestrian was crossing between cars (55.4%).

9.4 Discussion

The aim of this chapter was to understand how different pedestrian accidents occur in terms of different accident factors and accident causation scenarios with the other interacting road user. This analysis was used to discuss possible implications for countermeasure development.

9.4.1 Descriptive analysis

The descriptive analysis of the accident data showed that the pedestrian was the main contributing road user to the accident in 73% of the accidents in this dataset, and the road user did not expect the pedestrian to perform their behaviour in over 50% of all accidents coded. These findings are similar to the study carried out on urban pedestrian accidents by Carsten et al. (1989) where two pedestrian groups were analysed and adult pedestrians (n=297) were determined at fault in 71% of the accidents and child pedestrians (n=166) were determined at fault in 80% of accidents.

For the two largest groups of pedestrian failures detection and overall failures 0-13 year old pedestrians accounted for close to 50% for detection failures and over 50% of overall failures. Both of these groups had pedestrians as the primary contributing road user in over 95% of the cases. With regards to the cluster where the other road user was the primary contributory user pedestrians were not contributory to the accident occurrence in over 95% of the cases. For these cases 18-29 year olds accounted for close to one third of them, while they accounted for close to one fifth of the cases in total.

Close to 70% of the accidents occurred in urban areas and during the daytime, and close to 90% of the cases occurred in areas with speed limits of 30 mph or less. These figures are similar to UK national and European pedestrian accident figures (DfT, 2007; SafetyNet, 2009). The nature of the traffic environment and pedestrian facilities within it makes it realistic that most pedestrian behaviour will occur in urban areas and so these figures are to be expected. Close to 80% of the accidents occurred not at intersections but on road sections with the pedestrian crossing the road at a non-designated pedestrian crossing or crossing between vehicles and as such making it difficult for the other road user to anticipate the pedestrians behaviours. These results are similar to the finding of two studies where 75% of UK STATS19 national data showing pedestrian accidents occur where there is no crossing (Alnaqbi, 2009) and an analysis of Israel pedestrian accidents where 80% of accidents occurred on arterial multi lane streets (Gitelman, Balasha, Carmel, Hendel, & Pesahov, 2012). The high percentage of these accidents further underline that in close to 75% of the accidents the pedestrian was identified as the main contributing road user. In intersections, pedestrian crossing areas and areas where the other road user needs to make decisions, pedestrians can be expected to use more thorough visual search patterns.

The road types were the accidents occurred were most commonly either A roads or Minor roads, and with regards to injury accidents the accidents in A roads had a higher rate of pedestrian fatalities and serious injuries compared to other road types. This was possibly due to the higher speed limits in these roads as the relationship between vehicle speed and pedestrian injury mechanisms is well documented (Richards, 2010; Zegeer & Bushell, 2012).

With regards to the data the groups that are large enough to draw possible interpretations in terms of their failure mechanisms are overall failures (n=76), detection failures (n= 67), prognosis failures (n=57) and decision failures (n=49).

With regards to injury mechanisms the second highest percentage of fatal injuries (10.5%) and the second highest percentage number of pedestrian serious injuries (40.4%) occurred in the prognosis grouping where the pedestrian was not contributing to the accident (96.5%) occurring, and the other road user was either in a hurry or risk taking (table 73). The other failure mechanisms that had a high rate of fatal injuries were decision failures (14.3%) where the pedestrian undertook a risky behaviour and was either in a hurry, disobeyed a traffic sign or displayed unexpected/eccentric behaviours, and overall failures (7.9%) where two groups of pedestrians can be observed; The first were young pedestrians playing in the road side and not understanding the situation that they are in and the second were pedestrians that were intoxicated and failing to process the environment around them adequately.

9.4.2 Cluster analysis

In the cluster analysis four clusters were identified with three clusters relating to cases where the pedestrian and one cluster relating to cases where the other road user was the primary contributing road user to the accident. From these cases three distinct groups of pedestrian accidents can be determined.

Cluster 1

The first cluster identified a situation where the pedestrian was crossing the road either in view or from between vehicles where the other road user was going straight or at a traffic light intersection. The pedestrian's main reason of failure was detection issues (66.2%) due to their being in a hurry and not adequately scanning the roadside for other vehicles. The other road user was not expecting the pedestrian to make a manoeuvre and in just over half of the cases did not make an emergency reaction. The pedestrians age was evenly spread throughout most of the age groups, as were the other road users.

Cluster 2

The second cluster was the only grouping where the other road user was the main contributing user to the accidents occurrence. The accident occurred either due to two other road users colliding and then colliding with the pedestrian or due to the manoeuvre that the other road user was making. The pedestrians in this cluster were significantly older compared to the other clusters and the conflicting road user was significantly younger. The other road user did not make an emergency reaction for this cluster,

Cluster 3

The third cluster identified accidents that occurred during the night where the pedestrian either undertook risky behaviours or consumed alcohol which led to an accident occurring due to making a risky decision and not anticipating the risks of that decision. These accidents occurred when the pedestrian was either crossing the road or an intersection. This cluster is in line with findings in a number of studies about intoxicated pedestrians (Bradbury, 1991; Carole Millar Research, 1998).

Cluster 4

The fourth cluster highlighted an accident where pedestrians that were of a younger age, due to either being in a hurry or as a result of playing, confronted the road user in such a way as to cause an accident occurring. The pedestrian came out from in between cars to the roadside and the road user did not expect such behaviour to occur. In both of the situations the other road user reacted with an emergency manoeuvre more often than not but still could not stop the accident from occurring.

Cluster analysis general discussion

The cluster that identified the accidents that occurred where the road user was the main contributing vehicle had a high incidence of vehicle to vehicle accidents causing a pedestrian accident. The first accident type occurred in intersections and more complex road settings within a road that had an under 30 mph speed limit. In close to 70% of these cases the other road user did not make a mitigating manoeuvre. In this cluster the primary contributing road user was of a younger age. This can be a reason of the road users not performing an adequate scanning of the environment in a number of these cases. Borowsky, Oron-Gilad, Meir and Parmet (2012) presented novice and experienced road users with a task of tracking pedestrians in urban and residential areas and found that experienced road users processed information more efficiently than young-inexperienced road users (both trained and untrained) when pedestrians were identified.

In accidents where the pedestrian was the main contributing road user to the accident the other road user more often than not attempted to avoid the accident (either braking or going to the right or left). All of the cluster groupings where the pedestrian was the main contributory road user demonstrated that when the pedestrians cognitive functioning is affected by either impairment or being in a hurry/playing the pedestrians reactions are unpredictable for the other road user.

9.4.3 Countermeasure indications

When analysing the pedestrian data the pedestrian can be seen as the main contributing road user to 73% of the accidents in this dataset. Of the four clusters identified only one cluster had the other road user as the primary contributing road user to the accident occurring. This demonstrates that pedestrians performing unexpected behaviours or rule breaking lead to a majority of these accidents occurring.

For young adult pedestrians a cluster defined alcohol consumption and/or risk taking as one of the problems. Previous research in Scotland demonstrated that nearly a third (31%) of all pedestrian casualties had consumed alcohol prior to their accident compared to 5% of drivers and 9% of car passengers. These pedestrians were more likely to be males and between the age range of 20–29 (Bradbury, 1991; Carole Millar Research, 1998).

This is a difficult group of accidents to provide countermeasures to, as there are no laws to prevent pedestrians from walking while intoxicated or stop them from performing risk taking behaviours. Education and information with regards to alcohol consumption and its effect on individual motor control has been freely provided over the years particularly to road users and also to young adult pedestrians, advising them to stay with friends and not walk alone. For younger pedestrians a particular problem was their coming out in between other vehicles, as other road users had a problem reacting to this situation as their expectations were that this would not occur. It is difficult to expect young children to adhere to similar road safety standards as adults, so measures to restrict these types of interaction areas are necessary. A number of measures that were discussed during the previous decades have been implemented such as (Davies, 1999);

- 20 mph roads
- Traffic calming
- Speed enforcement cameras
- Publicity campaigns

Though these measures don't tackle the immediate issue of pedestrians behaving in an eccentric/unexpected manner on the roadway, they aim to lessen the consequences of the conflict situation and provide both pedestrians and other road users with a way of responding to conflict situations with mitigating behaviours. Changing pedestrian behaviour is something to be considered over the long haul and as an ongoing battle to provide better facilities and better education when using these facilities.

In cases where the other road user was the active component in the accident the main cause of accidents were the road user being in a hurry, taking risk or visibility issues in relation to the road environment. Visibility issues for the driver were present in 30% of the cases analysed and are one of the main issues that need to be tackled. Active safety systems that could help in directing the road users gaze into the direction of an event occurring could be particularly useful for better emergency reaction measures.

Focusing on the other road user to be particularly aware of these situations and providing training to them can also be another way of preventing these situations, but the cost benefits for these behaviours may not be high. The pedestrian conflict situations identified were similar in nature to those identified by Habibovic and Davidson (2012). Though due to the difference in the data collected they identified more factors related to the other road user's behaviour compared to this study, as it is difficult to identify road user's obstruction of view data from retrospective accident studies. The data provided highlighted that warning systems for other road users that are interacting with pedestrians are needed, the exact nature and timing of the warning in order for it to be able to affect the road user would be dependent on other factors that were not identified here, such as the exact pedestrian and vehicle trajectories during accident occurrences (Habibovic & Davidsson, 2011).

9.5 Summary

Chapter 9 described a statistical analysis of all of the Pedestrian accidents collected during the OTS data with regards to human failure and different accident factors. The study analysed 265 cases from the OTS dataset for the descriptive analysis and 248 cases for the latent class cluster analysis.

The results identified 4 accident scenarios for pedestrians, identifying three accident types where the pedestrian was the primary contributing road user and one accident scenario where the other road user was the primary contributing road user. Results indicated that in the majority of cases the pedestrian's behaviour initiated the accident situation and countermeasure indications on how to mitigate the accident situations with regards to the accident configurations focused on technological advances and pedestrian behaviour adaptation through training, environmental countermeasures and education were discussed.

10 General Discussion

10.1 Introduction

Understanding the causes of road accidents requires an understanding of the road, vehicle and road user factors that generate each individual crash. Such factors may interact, influencing the behaviour of an active participant in a crash just as the behaviour factors of one participant may interact with another. The main aim of this thesis was to identify prominent accident scenarios leading to traffic accidents employing a methodology able to incorporate these interactions. Development of a methodology to understand which factors interact with each other when an accident occurs, allows for both deeper interpretations of the factors and for countermeasure implications with regards to the transport system for these factors interactions to be carried out. This type of development provides a deeper classification of how errors occur compared to an understanding of individual risk factors.

Specific objectives were to (1) develop an analysis method that would allow for statistical analysis to be carried out to develop causation sequence chains in large traffic accident dataset, (2) analyse all relevant accident scenarios in the UK On the Spot accident study (OTS) dataset, (3) analyse the causal chains to understand how functional failure sequences occur within particular accident groups to develop accident scenarios, (4) understand the links between interacting factors and individuals to further understand how these interactions cause accidents to occur and (5) identify countermeasure implications for the scenarios that are highlighted in the research with regards to different stakeholders in the road environment. These objectives were made in order to first identify a suitable statistical procedure, demonstrate how it works with real world traffic accident data and develop accident scenario sequences to allow for a discussion of countermeasures.

The analysis carried out in this thesis aimed to demonstrate a method that would allow for all relevant accident causation and human error factors to be linked in terms of their interaction to each other and their interaction with those of other active road users. The reason for this analysis was to use a methodology that would identify accident causation analysis factors combined with other important accident site and vehicular factors. This would in turn allow a more detailed understanding of human error, compared to understanding accident factors individually. This research investigated the benefits of using a latent class clustering (LCC) model with real world in-depth accident data combined with accident causation data to understand accident scenario sequences.

The research comprised of one literature review, a method comparison study and four empirical studies focused on identifying functional failure sequences in traffic accidents. In this section the implications and contributions of the empirical studies with regards to the aims of the thesis will be discussed. Four separate studies were carried out, these studies focused on:

(1) Key accident scenarios in single and multiple vehicle accidents to demonstrate accident scenarios based on microscopic accident data.
(2) Comparison of national and in-depth accident scenarios to identify the differences between macroscopic (national) and microscopic (in-depth) accident data in terms of both level of detail and cluster results.
(3) Common accident patterns in powered two wheeler (PTW) accidents to identify common accident sequences for single and multiple vehicle accidents that include PTWs.
(4) Factors that lead to accidents with pedestrians in the traffic environment.

This chapter presents a discussion of the rationale, development and main results that emerged from this research under the following areas.

- Findings of the in-depth accident data analysis studies.
- The development of the method.
- Challenges and limitations of the applied methodology.

10.1.1 Overview of the work

Accident scenarios are clusters of accidents with a similar set of crash characteristics, the analysis of these accidents allows for identification of pre-accident and accident factors that increase the risk of accidents occurring. The analysis of accident scenarios can be used to develop prevention strategies to chains of events (Fleury & Brenac, 2001). This in turn allows for links between possible countermeasure implications to be made based on the results of the accident scenario development procedure.

The methodology has been used by IFSTTAR in France where in-depth accident (microscopic) cases are used to develop accident scenarios and national accident (macroscopic) cases are mapped onto these scenarios where possible, allowing for countermeasure implications to be made (Van Elslande, 2000) and in Sweden using the FICA dataset where accidents that have similar configuration such as intersection accidents causation charts are aggregated to see similarities and patterns (Sandin, 2009).

In order to be able to demonstrate the usability of the latent class clustering method real world in-depth accident data collected from the OTS accident study between the years 2000–2010 was used. The OTS study is based on stratified random sampling procedure.

Each study that used the OTS data had two sections, a descriptive analysis and a latent class cluster analysis. To illustrate the different factors and accident groupings the results of a descriptive analysis were reported within each chapter. This analysis provided a snapshot of the data, and helped the interpretation of the cluster analysis by providing an overall picture within the specific datasets. This analysis allowed an interpretation of how each individual factor was present in the overall accident data. The latent class cluster analysis was carried out to link variables that interact in specific clusters. A chi square analysis with an analysis of the residuals in each of the clusters was also carried out in order to identify significant values within the clusters. The analysis of the residual values for the categories within each variable allowed for a differentiation between significant and non-significant values to be carried out. This allowed for cases where a variable value was over-represented to be brought to the forefront of the analysis and compared with different scenarios as well as the overall dataset.

For example in the single vehicle cluster analysis the road user's gender was male in over 70% of the accidents. The chi square analysis allowed a further interpretation of the clusters in which males (clusters 1 and 6) and females (cluster 2) were over-represented compared to the overall values. The descriptive analysis results were also taken into consideration so the chi square analysis was an extra analysis step rather than a separate analysis method. The cluster analysis results were then interpreted as accident scenarios and discussed in terms of how the significant variables interacted with each other.

One of the difficulties when conducting this work was to connect it to previous traffic accident research. One aspect of this approach was the use of the perceptual Human Functioncal Failure (HFF) causation methods failure model as this was not commonly used in other research using cluster analysis (de Oña, López, Mujalli, & Calvo, 2013a; Depaire et al., 2008; Skyving et al., 2009), and where an accident causation coding method was used with a descriptive analysis, only a very small number of cases were analysed (Ljung Aust, 2010; Sandin & Ljung, 2007; Van Elslande, 2000). Nevertheless these connections were made in the discussion sections of the studies.

10.2 Findings of the research

10.2.1 Analytic results

The detailed review of previous research found that the highest level of detail for an individual accident is provided by in-depth accident data collection methods. The main advantages of this method are the detailed information that is provided with regards to factors related to the roadside/infrastructure, the vehicle and the road user. Interviews conducted with active participants provide detailed information on the road users actions before the accident and questionnaire data that is obtained following up the incident provides further details.

Despite providing the highest level of detail due to low case numbers in-depth accident causation studies have not previously used any statistical methodologies to group accidents together based on similarities.

A total of 25 scenarios were identified by the latent class cluster analysis in four separate studies using OTS data (single vehicle accidents, multiple vehicle accidents, powered two wheeler accidents and pedestrian accidents). A number of these clusters had overlap in terms of the human functional failure perceptual stages that they were alluding to, though the different infrastructure and vehicular factors that were present in the scenarios allowed for countermeasure implications to be made with regards to similar accident groupings.

Accident dataset population study

All single vehicle accidents and multiple vehicle accidents collected between the years 2000-2003 were selected from the full OTS dataset. A total of 1,614 accident cases were collected between these years, for the cluster analysis only cases that included all relevant accident data variables were selected. Any case that did not include information on variables that were included in the cluster analysis were omitted. For the latent class cluster analysis 366 single vehicle accidents and 673 multi-vehicle accidents were used.

The descriptive analysis carried out on the OTS data identified that a majority of the cases had detection failures and prognosis failures as the main human functional failure types. The most common types of accidents were lane changing, rear-end and loss of control accidents. The most commonly coded human factors were speeding, the road user being in a hurry and the road user breaking the law. Accidents occurred mainly in lower speed (30 mph) or higher speed (60-70 mph) roads. Over 75% of the vehicles involved in the accidents were cars.

Amongst the single vehicle accidents several scenarios were identified. Four of the scenarios occurred on high speed roads and two occurred on low speed environments. The three largest accident groupings were different types of leaving lane accidents, the largest group related to road users making diagnosis errors due to speeding, the second due to road users making a detection error due to being in a hurry or the road condition and the third due to alcohol and impairment and the road user making a decision or execution error. A comparison with an analysis of Swedish data (Sandin & Ljung, 2007) showed a similar pattern for the more common scenario but lower commonality for the smaller groups of accident scenarios. The clusters that had higher accident severity outcomes were the first and third group of clusters.

The latent class cluster analysis identified four clusters related to detection issues in multiple vehicle crashes, the road user not identifying the other road user. Two of the clusters highlighted issues with the driver's search patterns while turning right at a junction. These results are similar to findings for studies looking at 'looked but did not see' accidents (Brown, 2002; Clabaux et al., 2012; Clarke et al., 2004; Clarke et al., 2007; Crundall et al., 2012; Hurt et al., 1981). The first of these two clusters grouped together 'looked but did not see' accidents at intersections involving both car to car and car to PTW accidents. This result indicated that similar types of interactions occur with regards to visual search patterns for these accidents despite the different vehicle types. These clusters indicated that the road user did not use appropriate visual search patterns due to either being in a hurry or their emotional state. One cluster identified rear-end accidents where the driver did not detect the road user ahead. This grouping contained most rear-end accidents within the dataset with a majority being on higher speed (60-70 mph) roads. The last cluster highlighted situations where the road user was overtaking a vehicle on a motorway and did not recognize that the situation was not suitable for an overtaking manoeuvre and has poor control of their vehicle. These groupings are in line with the most common type of multiple vehicle accidents in the STATS19 dataset being rear-end accidents, overtaking accidents intersection accidents (DfT, 2013), however this analysis provided new insight with regards to the accident setting and human errors for the accident types.

Powered two wheeler study

All powered two wheeler accidents collected between the years 2000-2010 in the OTS dataset were used. A total of 339 accidents were included in the descriptive analysis for multiple vehicle accidents that included PTWs and 428 accidents including single PTW and other vehicle to PTW accidents were included in the cluster analysis. The selection for the descriptive analysis was made to identify important factors when the PTW was in an accident with another road user. All PTW accidents were included in the cluster analysis to identify similarities and differences for these accidents.

The descriptive analysis identified that for the PTW rider involved in two vehicle accidents the most common human functional failure was a prognosis failure. This demonstrated that the other road user was making a behaviour that was unexpected by the PTW rider. The main contributory factors were the rider being in a hurry and speeding. Most PTWs were larger than 250cc in engine capacity and the other road user did not make an emergency manoeuvre. PTW riders were considered the main contributing road user to the accident in only 36% of the accidents, which was lower than findings of 51% from a study by Clarke et al. (2007).

The cluster analysis separated single PTW accidents and multiple vehicle accidents succinctly. ‘Looked but did not see’ accidents where road users did not identify the PTW rider were found in two clusters. This is similar to other studies in the area (Clarke, Ward, Bartle, & Truman, 2004; Clarke et al., 2007; Crundall, Crundall, Clarke, & Shahar, 2012). For single PTW accidents the rider was considered as taking unnecessary risks related to speeding in regard to the different situations confronted. This is in line with self-reports from PTW riders in terms of speeding in a majority of situations (Clarke et al., 2004).

Pedestrian study

All pedestrian accidents collected between the years 2000-2010 in the OTS dataset were used. A total of 265 pedestrian accidents were used in the descriptive analysis and 248 accident cases that included all relevant data were used in the latent class cluster analysis.

The descriptive analysis identified that the contributory factors for pedestrians were most commonly being in a hurry, a young pedestrian playing and alcohol. The common human functional failure types were overall failures and detection failures. Pedestrians were considered the main contributing road user in 73% of the accidents, meaning that the accident investigator considered the pedestrians behaviour as the main initiator of the accident. These findings are in agreement with a study carried out by Carsten, Tight, Southwell, & Plows (1989). Close to 80% of the pedestrian accidents occurred at non-designated pedestrian crossing road sections which is consistent with STATS19 data (Alnaqbi, 2009) and a study carried out on pedestrian accidents in Israel (Gitelman, Balasha, Carmel, Hendel, & Pesahov, 2012).

The cluster analysis expanded on the descriptive findings by identifying four main pedestrian accident scenarios. The first occurred when pedestrians did not detect other vehicles in the traffic environment due to being under stress or limited visibility. The second was due to the other road user not detecting the pedestrian due to being in a hurry. The third was a night time accident where the pedestrian was impaired or had taken alcohol and made a violation or lost cognitive capacity. The fourth involved younger pedestrians crossing from between vehicles and not being seen by the other road user.

STATS19 study

In order to highlight the different results that are obtained from microscopic and macroscopic accident data a latent class cluster analysis of all national data collected in Great Britain for two vehicle accidents from the year 2004 was carried out. A total of 55,474 accident were included in this analysis,

The cluster analysis identified large clusters that were based on accidents occurring in a T or staggered junction in a give way setting in a lower speed area. Four of the clusters were related to single carriageway accidents, and the differences between the clusters were based on the road users manoeuvres in relation to one another. Three of the clusters were nearly identical in nature relating scenarios where one road user turned and another road user was going ahead on a T junction. The similarities of the results made it difficult to provide a relevant discussion in terms of countermeasure implications. The results found in this analysis were similar to other latent class cluster analysis studies that also used microscopic data (Depaire et al, 2008; De Ona et al., 2013), in that the cluster groupings were characterised either by traffic accident characteristics or road user contributory factors.

The comparison of microscopic (in-depth) accident data using the UK On the Spot dataset and macroscopic (national) accident data using the STATS19 dataset demonstrated that in order for detailed clusters that include relevant human, vehicular and infrastructure/environmental factors to be formed, data that includes accident causation information is needed. This type of data is only provided by in-depth accident data. When the variables and variable values that are coded by the accident investigator are limited with regards to the accident coding of one aspect of the accident, the results of the cluster analysis are skewed towards the variables that allow for a clear differentiation between variable values.

In the analysis the road user 1 contributory factor was most commonly clearly significant for the value error or reaction and there was little differentiation for the clusters whereas the junction type variable was separated based on the variable values. This was particularly an issue for road user based human factors that were mainly limited to demographic factors that did not allow for a clear interpretation of the resulting clusters. For example gender and age ranges were significantly over-represented in the resulting clusters but this was due to the large case numbers, the percentages were mostly within 5% of the overall percentages for the variable values.

10.2.2 Methodological findings

Accident scenario analysis using the Human Functional Failure method

The studies carried out in this thesis applied the Human Functional Failure (HFF) method that was developed and used by IFSTTAR in France. This methodology has been previously used in countries other than France, within the Traffic Accident Causation in European project (Naing, Bayer, Van Elslande, & Fouquet, 2007; Van Elslande & Fouquet, 2007) and in the Netherlands (Boele-Vos et al., 2016). Other than a small number of cases coded in the TRACE study (Naing et al., 2007) within the UK, this was the first time that this large a number of cases were analysed and coded in a UK sample. The studies carried out using HFF demonstrated that identification of general human error and accident factors translates over to UK data, though the accident scenarios need to be developed individually.

A number of advantages were observed with regards to using the Human Functional Failure method compared to previous LCC analysis studies that did not use accident causation coding systems (de Oña et al., 2013b; Depaire et al., 2008). In the study carried out by de Oña et al. (2013b) variables related to the road user that were used in the analysis were gender, age and accident causes. The accident causation variables were grouped into one variable that differentiated causes into four possible categories; driver characteristics, road characteristics, vehicle characteristics and other. Driver characteristics were most commonly identified as being the main cause of the accident (92%) by the accident investigators. In the study carried out by Depaire et al. (2008) variables related to the road user were gender, age, and behaviour. The behaviour variable was grouped into ignores red light, fails to give right of way, crosses a full white line, passes incorrectly, makes an evasive manoeuvre, incorrect location on the road, loss of control, not enough distance kept, fall and normal behaviour. The categories are concentrated on issues related to road user violations and are limited in the explanation of the road users selected behaviours.

The analysis of in-depth accident data coded with HFF carried out in this thesis allowed for a deeper differentiation of human factors compared to the other studies that used cluster analysis. The human factors were divided into distinct groups that illustrated differences between violations, lapses in attention and error based failures. In addition to this coded information two contributory factors were included for the single vehicle accident analysis, and one factor for each road user was included in the multiple vehicle, PTW and pedestrian accident analyses. These factors were included in addition to road user’s gender and age. This differentiation allowed for the road user behaviours leading up to the collision and other factors to be coded together, and for the statistical analysis to contain a larger number of human related factors that have been collected based on an human functional failure coding methodology. The inclusion of the main human functional failures allowed for an understanding of the manner in which the pre-crash infrastructure and vehicular characteristics together with traffic conditions in combination resulted in specific errors of one or more road users.

Historically accident studies have focused on understanding road user error within the traffic environment (Carsten et al., 1989; Sabey & Staughton, 1975; Treat et al., 1979). However the data did not include higher level information on the operation of the transport system and it was concluded that if further data became available in future studies than the analysis would be further enhanced. This methodology would require all relevant factors to be identified and related to the higher level system based information such as legislation, stakeholders in the system and the functioning parts of the system to be merged together and interpreted as a whole. An example of the work that would be required can be found in an article by Salmon, Read and Stevens (2016) where a detailed structural analysis of the working structure of the road transport system operation in Queensland, Australia was carried out.

Some of the limitations of using the HFF method relate to the Safe System approach. While a large number of factors to be coded in terms of the traffic system are present and can be coded with the HFF method, the system as a whole has not been directly modelled, and as such the interpretations that the analyst makes with the crashes may be limited in terms of higher level system based factors (design, operation and management issues). A merging of this method with a more system based approach such as those used by (Reason, 1990; Reason, 1997), could help bridge the gap between in-depth accident studies and methods used as a systems approach.

Statistical analysis methodology

The preparation and analysis of the data was a long procedure that included a number of false starts with regards to the data analysis that was selected. A comparison of different statistical methods that could have been incorporated into this thesis was carried out during the data preparation stage. From this review, a number of methods were identified as being suitable for use with the data when considering the analysis requirements. Two of the methods that were identified as possibly suitable, principal component analysis and quasi induced exposure methods, were tried.

These two methods were found to be unsuitable due to their inability to compartmentalize the analysis allowing for a comparison of all interacting factors. The principal component analysis did not clearly identify accident scenarios and grouped most of the factors within a small number of scenarios. This limited the ability to clearly identify interacting factors with a number of scenarios, and also caused a loss of data within the analysis. The quasi induced exposure method required a control group to be selected from road users that were identified as non-active within each accident. This limited the interpretation of the interactions of each individual accident scenario, as a comparison of factors was going to be necessary. Regression methods were also considered but ultimately discarded as they required one or more outcome variables for interpretations purposes.

Ultimately the latent class clustering method was selected for this thesis as it was determined that this method was most applicable to the purpose of this research to develop scenarios without concentrating on outcome variables and allowing each variable to have an individual weighting with regards to the overall scenarios. One of the main advantages of this methodology was that the analysis did not lose any of the information when identifying variable interactions and included all variables within the analysis results, whether the individual findings were significant or not.

The statistical measures based on data mining used in this thesis can be used as a bridge between quantitative and qualitative methods in analysing accident data to uncover hidden relationships, as historically most analysis methods concentrated on dependent factors to analyse the varying level of risk. The aim in this thesis was to not concentrate on injury risk, but rather to identify where failures occur and so LCC was the most suitable method for this purpose. LCC analysis allows a definition of risk while relating multiple factors together within a cluster.

The interpretation of accident data requires a consideration of the research question and then the accident data. When modelling accident data using statistical models, it is necessary to consider whether the variables within the data have linear properties. Modelling linearity is particularly necessary when using parametric based modelling methods such as hierarchical cluster analysis and hierarchical regression. The assumptions of linearity require the data to include an outcome variable, which in traffic safety studies is most commonly risk.

When considering the suitability of the statistical method, two different considerations were made. Firstly the advantages compared to previous road safety studies using LCC, and secondly the suitability of the method compared to other possible statistical procedures.

The measurement of a large amount of human related factors in traffic crashes requires a subjective assessment by using interview data and accident reconstruction, as it is not sufficient to collect information based on the accidents physical attributes with regards to human error. These factors are then combined with physical factors that are either binary (yes/no) in nature or to measurements related to the road infrastructure/presence of environmental factors. In the analysis chapters, objectively and subjectively measured accident factors were combined in such a way as to allow for direct statistical analysis to be carried out, and as the analysis was concentrated on identifying accident types rather than outcome variables such as injury information, the purpose was to develop scenarios of interacting factors. The highest possible level of detail was provided on both the descriptive and cluster analysis level, and the analysis concentrated on making sure that the statistical method did not allow for any data to be lost during the analysis.

When considering accident data analysis techniques no one technique will provide all of the answers to all of the questions. It is necessary that different questions are answered by using different data sources and statistical procedures, for the issues that were examined in this thesis, the use of other statistical methodologies that are commonly used in accident data analysis, were not appropriate. The issues related to principal component analysis and quasi-induced exposure methodologies were stated above. The use of regression based analysis methods is dependent on the identification of one or more dependent variables, most commonly injury outcomes. The inclusion of a dependent variable would provide clarity with regards to that variable, in terms of the injury outcomes, but not allow for clear interpretations of interaction between different variables to be made.

When considering accident data, categorical and continuous data need to be analysed together. This requirement makes it difficult to include all relevant data within models that provide a basis through a linear relationship. If an analysis of accident risk with regards to injury outcomes was carried out there are more suitable methods for risk analysis than LCC, such as neural networks (de Oña et al., 2013b) , multinomial logit estimates (de Oña, López, & Abellán, 2013) and logistic regression procedures (Michalaki et al., 2015). These types of analyses provide detailed numerical information in terms of risk with regards to injury outcomes. An advantage of using the LCC method compared to other multivariate methodologies is that LCC allows for the interaction of the variables to be mapped together rather than observing individual risk outcomes or having to group variables together for this purpose. Decision trees (Badea – Romero & Lenard, 2013) and other regression based tree procedures (de Oña et al., 2013) are similar in nature, though cluster analysis provides a clearer allocation of all variables within the clusters by determining the variables by a percentage rather than an 'either-or' selection procedure.

The use of these methods compared to the descriptive based approach allows for relationships that were not previously detected to be observed. This helps bring a level of objectivity to the understanding of accidents rather than grouping accident types according to a logical structure. The grouping of accident studies using expert judgment based on a logical framework for comparison either by accident configuration or human failure type has been used by a number of studies. The Factors Influencing the Causation of Accident and Incidents (FICA), the German In-Depth Accident Study (GIDAS) & Institut National de Recherche sur les Transports et leur Securité (INRETS) studies have been commonly used for ADAS advancement purposes. The method used in this thesis helps provide a comparison point to previous studies that have aimed to use risk factors or other statistical measures for macroscopic accident analysis purposes.

The limitations of working with a basic system of coding accident causation was also demonstrated in the national data analysis chapter where although accident causation variables were coded, a clear linkage to other factors was not possible due to the nature of the codes.

The division of the powered two wheeler (PTW) and pedestrian accident cases into separate analyses furthermore allowed for the identification of specific accident types for these road user groups rather than concentrating solely on the crash configuration of pedestrian/PTW against other road users. This is particularly pertinent when considering that in the accident analysis chapter all of the pedestrian accidents were grouped in two large groups that did not allow for a differentiation of the data. A similar issue was seen in the study by Depaire et al. (2008) where all pedestrian accidents were grouped into one large cluster. In both the present study and the study carried out by Depaire et al. (2008) the accidents involving a pedestrian were considered by the cluster analysis to be significantly different from the other accident types. This resulted in the clusters only including pedestrian to other vehicle accidents and not allowing for a differentiation of different types of pedestrian accidents. The analysis carried out on the vulnerable road user groups allowed for individual clusters to be identified.

The segmentation of the accident groupings, particularly with microscopic data that provides clearer definitions compared to previous analysis, allows for specific groupings to be analysed. This can be seen in the differentiation of the ‘looked but did not see’ accident groupings where three different accident types were identified. The inclusion of errors made by the drivers of the interacting vehicle also allowed the other road user’s expectations and reactions to be considered. This type of understanding cannot be obtained from macroscopic data without information on road user error. If the number of in-depth accidents collected could be increased, the possibilities of the cluster analysis would also increase proportionally.

Compared to previous accident causation based accident studies which analysed 38 single vehicle crashes (Sandin & Ljung, 2007) and 392 road users involved in crashes (Van Elslande, 2000), this thesis had a significantly larger number of cases analysed and compared with an in-depth statistical procedure. The level of data available and the methods that were used to gather this data were the most advanced available at the time.

The FICA and IFSTTAR studies have been used by both road operators to help implement changes in the traffic environment and by vehicle manufacturers to help develop active safety systems that will address the issues of road users in the traffic environment as identified by analysis of accident causation data. The Volvo Car Corporation and SAAB were partners in the FICA study. The HFF, ACASS and DREAM models have also been used in a number of different European traffic safety projects for accident analysis and countermeasure development purposes.

10.3 An examination of the results in relation to existing research

The literature review of models related to driving and human error provided a detailed description of how human failure can influence accidents and what type of failures can occur to cause a traffic accident. Previous research identifying failure in the case of traffic accidents has commonly quoted that human error causes up to 95% of all accidents (Treat et al., 1980; Sabey & Staughton, 1975). A detailed understanding of human factors within a road accident is necessary in order to understand how road users are interacting within the traffic environment and how human errors are initiated.

The results of the studies carried out in this thesis demonstrated that an analysis of a significant number of detailed factors within the accident are necessary to understand the accidents as a whole. Countermeasure development is increasingly based on a Safe System approach and when considering road user error it is necessary to take into account the other relevant higher level factors that lead to the road user making an error. Past research has focused on developing models to help demonstrate how human failure (Rasmussen, 1982; Reason, 1999; Norman, 1981) occurs. Much of the research that has been carried out in in the past 10 years on accident causation has used these models or similar models as a basis for the development of tools to understand human accident causation (Ljung, 2007; Otte, Jaensch, & Pund, 2007; Van Elslande & Fouquet, 2007; Wallén Warner, Björklund, Johansson, Ljung, & Sandin, 2008). The advantage of using these types of models to understand traffic accidents is that accidents can be grouped in a relevant time based sequence and can be analysed in more detail for countermeasure purposes. These models also allow for human factors and other factors to be determined for each accident in a holistic approach. Nevertheless these models are typically applied to single accident cases and are only rarely used for aggregate data analysis.

The use of the above stated accident causation methods to classify human error makes it possible to determine how certain failure types and contributory factors contribute to an accident. Before deciding on the accident causation coding model to be used within the thesis, a comparison of three possible models was carried out. This comparison aimed to clarify the usability of the different models with regards to the aims of the studies.

The models compared were the Driver Reliability and Error Analysis Method (Ljung, 2007), the Accident Causation Analysis with Seven Steps (Otte et al., 2007) and the Human Functional Failure (Van Elslande & Fouquet, 2007). Ultimately the Human Functional Failure method (HFF), a method developed in France that was based on previous research in human failure by Rasmussen (1982) and Reason (1990), was selected to be used. The main advantage observed in the HFF method compared to the other two models were that clearer interpretations of all factors related to the collision occurring compared were possible. Despite this advantage there was a disadvantage in terms in inter-rater reliability compared to the other two methods.

The HFF method uses a five stage perceptual model related to how road users perceive the traffic environment. The road user is constantly going through these stages when making decisions related to their behaviours. The road user first perceives (stage 1) the information from the environment, then diagnoses (stage 2) the situation, anticipates (stage 3) how events will unfold, makes a decision (stage 4) and then performs an action (stage 5). A sixth stage related to issues with sensory or cognitive impairment/failures is also included. The HFF model had a detailed number of 30 top scenarios with regards to the above main failure that occurred for the road user based on French macroscopic and microscopic data. These scenarios can be seen in the Appendix in page 374. Current research analysing traffic accidents has focused on the detection and decision stages (MAIDS, 2009; Pai & Saleh, 2008; Pai, 2009) to distinguish when the road user makes an error, though the other stages have not been used particularly when macroscopic data is analysed.

This method allowed an analysis of the accident data with regards to accident causation measures as well as an analysis of the main failures that all road users involved in the accident made. When applying HFF, all relevant factors leading up to the incident with regards to the road user, vehicle and environment/infrastructure are collected using a mixture of objective and subjective data collection measures for coding purposes.

For the study carried out on multiple vehicles using the UK On the Spot (OTS) dataset, the coding method which further partitioned the six main failure groupings into 30 sub-groupings was originally used. This analysis was ultimately discarded because the cluster groupings did not provide relevant results, most likely due to having been developed from French data. The analysis based on the six main failures was carried out instead.

There were two main constraints with regards to using this expanded methodology, the first of which is theoretical and the second statistical. Firstly, these scenarios did not provide a broad understanding of the accident clusters, rather the resulting segmentation was not clear and provided a large amount of overlap between the cluster findings. A possible reason for this is that the nature of the UK driving environment is different than in France so although the broader categories were suitable for analysis purposes, the sub-categories being developed in France did not clearly represent the UK in-depth data. Secondly the number of cases within the accident dataset only allowed a limited number of factor values to be entered into the analysis. The number of factors that can be included in a cluster analysis are based on the number of cases that are present in the dataset. The cluster analysis studies carried out by Depaire et al. (2008) used 29 variables with 4,028 cases from a national dataset, de Oña et al. (2013b) used 18 variables with 3,229 cases from a national dataset, and Skyving et al. (2009) used 167 cases with 12 variables from an in-depth accident dataset. Approximately at least 10 cases per variable factor are accepted within accident data analysis for in-depth accident data and this was taken into account for the in-depth accident studies cluster analysis carried out in this thesis.

When using latent class clustering methods only a finite number of variable values can be entered into the analysis. For the main functional failure rather than entering a total of 30 factors into the analysis for the main failures, entering 6 main failure factors allowed for a more evenly distributed cluster analysis between all of the factors. This allowed for other factors related to the infrastructure, vehicular and environmental factor, such as the road lighting and road area type, to be also included in the cluster analysis, and the scenarios to be evenly weighted with regards to the different coded factors. This was done in order to have a better balance in terms of coded factor types compared to previous research (de Oña et al., 2013b; Depaire et al., 2008).

Previous work using accident causation data has concentrated on either case by case analysis methods (Ljung Aust, 2010; Sandin & Ljung, 2007) or subjectively relating accident types to national data and creating groupings according to this analysis (Van Elslande, 2000). These types of analysis allow for an understanding of how each individual accident was caused, but do not allow a determination of which factors are significant or over-represented in individual situations using statistical modelling measures. The advantages of using a method that is based on comparing factors using significance testing, is that when understanding the most commonly occurring accident scenarios, as in the present research, the weight of each individual factor and occurrence can be identified.

For the first road user within the multiple vehicle accident database and the single vehicle accident data, the stage of perception was similar to the results from Van Elslande (2000). The main stages of functional failure were detection failures and diagnosis failures. When the second road user was included in the analysis prognosis failures was the largest failure grouping.

The grouping of the two road users when analysing multiple vehicle accident data was unique in that it allowed an interpretation of both the accident setting and relevant human factors related to the accident together. This grouping allowed for further differentiations to be made with regards to the scenarios. For example in the PTW cluster 5 results both car drivers and PTW riders that made decision failures and prognosis failures were identified by the clusters. This indicated that both road users were making similar human functional failure types.

The analysis within this thesis, through the use of microscopic data, provided results that included greater detail compared to previous studies that used macroscopic accident data (de Oña et al., 2013b; Depaire et al., 2008). Each clusters countermeasure implications were discussed with regards to possible measures that could be taken, and clusters where similar countermeasures could be carried out were identified. The use of accident causation coding also allowed for identification of the accident configurations in great detail, identifying the factors in a time based fashion.

If considering the implications of the results, in Depaire et al.'s (2008) study using two vehicle macroscopic data in the Brussels region of Belgium the seven resulting clusters were titled (1) crossroad with no traffic light, (2) traffic accidents with adult pedestrians, (3) traffic accidents on crossroads with predominantly traffic lights, (4) traffic accidents between a car and a non-moving second road user, (5) traffic accidents with a motorcycle or bicycle, (6) traffic accidents with non-adult pedestrians and (7) traffic accidents on highways. A factor related to human behaviour was included in the analysis, but no relevant significant values related to this factor was reported within the cluster analysis results.

The results from the multiple vehicle accident scenarios study using OTS data carried out in this thesis were in order: (1) turning accidents in a low speed setting due to detection issues from visibility or lane violations, (2) rear-end accidents in a high speed setting due to detection issues, (3) urban road low speed accidents, (4) lane violation due to speed or impairment, (5) intersection accidents due to breaking the law, (6) right of way violations due to road user risk taking or illegal behaviour, (7) pedestrian accidents occurring as a result of impairment, and (8) pedestrian/cyclist to car accidents where the road user made the primary contributory behaviour.

The nature of the results differed in that specific accident segmentations and countermeasure implications based on these results were possible. This was not the case in Depaire et al.'s (2008) study where indications of accidents are present but a detailed understanding of the interactions leading to the incidents cannot be deduced from the cluster analysis.

The advantages compared to the previous study can be identified as: (1) the human factors were not present in the Depaire et al. (2008) study, while in the study carried out in this thesis they were, (2) The scenarios were more detailed in nature and a more detailed discussion of countermeasure implications were possible to be carried out, (3) A further analysis of the data using a separate statistical modelling approach was not necessary.

10.4 An indication of the importance of the findings

In-depth accident research studies have concentrated on using a model of human error types to provide individual analysis to accident cases and identify accident scenarios for analysis purposes, identifying Advanced Driver Assistance System (ADAS) measures as appropriate for individual accident configurations. These methods have commonly used descriptive statistics to help identify road safety issues either previously without the use of a behavioural causation model (Morris et al., 2006; Sabey & Staughton, 1975; Treat et al., 1979) or when using an accident causation coding model (FICA, GIDAS & ITS study).

The use of a statistical mining tool using microscopic accident data to help better understand factor interactions, as reported within this thesis, is unique as the numbers of accidents collected by researchers are commonly not large enough within in-depth accident datasets for this purpose. The higher number of cases collected within the UK On the Spot (OTS) accident dataset allowed a sufficient number of accident cases to be used for statistical modelling purposes. The main benefit of using this combined approach including in-depth accident data and the Human Functional Failure (HFF) accident causation analysis coding, were the more detailed results compared to previous studies that used national microscopic data and latent class clustering methods (Depaire et al, 2008; De Ona et al., 2013).

This methodology works in two ways, (1) it allows for an extension of previously identified accidents to be combined with a large number of relevant variables that give information about which situations accident scenarios occur in and (2) it allows for previously hidden or unknown relationships to be broached and detailed. The nature of in-depth accident data makes it necessary that clear methodological steps are necessary for analysis purposes, data mining approaches can help provide these possibilities.

10.4.1 Research and policy implications

When understanding accident data output it is necessary to put it into the big picture of how developing and developed countries are tackling issues related to traffic safety. Developed countries are using the Safe System approach based around measures developed within vision zero to decrease accident rates and fatal injuries suffered by road users.

When analysing in-depth accident data using causal inferences, most methods used a number of different crash causation classification systems, in order to uniformly identify accident factors. There has been difficulty in linking the factors identified by these models (human, vehicular & environmental/infrastructure) to other systems based information such as road rules and regulations (Salmon et al., 2010). The output that has been generated by using the HFF data interpretation method and LCC analysis has provided detailed information on how interactions of different factors within the road environment cause accidents, though a link to road rules and regulation would provide possibilities to make more detailed observations and interpretations.

The HFF method has been previously used in France to develop accident scenarios for analysis and countermeasure purposes. This is done by analysing data collected from in-depth accidents and identifying possible scenarios, and then mapping out this data onto national accident database data to identify similarities. The analysis carried out in this thesis has identified accident causation scenarios using a purely statistical approach that is different to the subjective analysis approach carried out at IFSTTAR. The use of an objective method removes the potential of the researcher's preconceptions framing the analytic outcomes.

The results of this analysis were clear cluster scenarios based on a large number of cases grouped together based on the main issues related to the individuals. These identified clusters could then be broken down into more specific groups for analysis purposes. Clusters that were similar in nature such as the leaving the lane accidents for single vehicles were analysed together, and indications for the different factors leading up to these incidents were separated based on the cluster results. For leaving lane accidents though the largest cluster was related to speeding the other two cluster results indicated that detection issues and impairment also led to similar accident occurrences. Incidents related to alcohol also more commonly occurred in rural areas.

In this thesis, the situations leading to driver detection errors were further elaborated on with regards to both the road user making the error and also the other road users that were involved. When considering accident data, it is better to consider all road user's interactions rather than each road user individually. Possible reactions from all of the road users that are interacting within the accident locus proximity should be taken into consideration, and future developments particularly with regards to active safety countermeasures need to work within the attention limitations that road users have. If a holistic approach considering all road users is used traffic safety measures could be made more reliable through the understanding of how individuals interact with each other within specific settings.

The structure of the traffic environment is dynamic in nature, and future developments should aim at eradicating all potential issues from the environment rather than concentrating on individual road user issues. A clear benefit of analysing each accident on a case by case basis and including all variables on this basis in the multiple vehicle accident analysis was that a clear scenario including both vehicle users could be developed. The inclusion of both variables that combine accident level variables for both road users, accident level variables for individual road users and human factor variables for each road user allowed for this data to be gathered in a way that was not done previously. This process allowed for a more detailed understanding of how interactions occurred in the accident configurations.

Similar to previous literature (Brown, 2002; Clabaux et al., 2012; Clarke et al., 2004; Clarke et al., 2007; Crundall et al., 2012; Hurt et al., 1981) ‘looked but did not see’ accidents, accidents related to inattention or related to distraction, were identified within the multiple vehicle and powered two wheeler accident analysis, particularly with regards to powered two wheelers (PTW). For these scenarios the inclusion of all relevant road users within the analysis allowed for a clearer identification of the interaction that was taking place between all road users.

The implications of the results for ‘looked but did not see’ clusters highlighted that, for the car driver PTW accidents, the incidents occurred on lower speed limit roads during the night-time and higher speed limit roads during the daytime. The lower speed limit scenario were commonly occurring on a junction whereas the daytime setting was on a motorway or on an A road and was related to the speed that the other road user was going at. These results elaborated previous results by demonstrating both the different accident situations that the incidents were occurring in as well as allowing a differentiation of these incident types and reactions by the other road users.

Most technological developments concentrate on alleviating issues with regards to the driver’s detection of the situation or of other road users in the traffic environment. But for all issues to be directly tackled all road users involved in the conflict situation should also be taken into consideration and future technological developments should aim at alleviating all sides of this issue, rather than just concentrating on the road user that is identified as the road user that is considered at fault for the accident.

The use of the accident causation method allows for an understanding of the human functional failure types and factors that cause different types of accidents, that helps underline the different human failure types that occur within the OTS dataset. Similar findings have been presented in previous research literature, but the interaction of these failures, with other types of contributory factors and accident level variables in this research is novel in nature. The inclusion of the errors of all road users within the cluster analysis also allowed for the contributory factors and main failures of both road users to be taken into account. The road environment setting of the accident, the type of accident occurring and the human factors interaction all allowed for a more thorough understanding of the cluster implications.

This thesis did not aim to analyse the scenario results with regards to the available vehicle technologies for active and passive safety, but the implications for this technology are still present. These implications were discussed with regards to the road user behaviour highlighted by the scenarios and mainly focused on countermeasures related to behaviour, though these could be elaborated with the codes on environmental/ infrastructural factors that were coded within the accident scenarios and individual cases. For example cluster 7 in the PTW accident analysis identified cases where the state of the road directly led to the accident occurring and as such the implications for these accidents are in terms of keeping the roads in a suitable manner for PTW riders, rather than an issue related to their riding behaviour.

10.5 Limitations of the study

The limitations of the whole thesis are discussed in this section within the general framework of accident causation. The main limitations of each of the studies are also discussed within this section. When taking all of the research studies into consideration a number of general limitations can be highlighted. The main limitations of the studies presented in this thesis can be summarised as:

- The human functional failure coding of the OTS data was carried out by one investigator retrospectively.
- Latent class clustering analysis interpretations are limited to identifying clusters but clusters had similar results and overlapped.
- An estimation of risk using another statistical method was not carried out.
- Higher level system operation data was not available.

The OTS data was collected by a team of individuals and coded after the accident and reviewed in the office once they returned. The accident causation data was coded by the author by retrospectively reviewing the available information available from the OTS accident database. Ideally each individual case would be coded by a team of accident researchers and discussed before coding is completed (similar to the procedure at IFSTTAR), though the time limitations and availability of other individuals to code this data was not possible.

This would limit the amount of subjective interpretation bias for each individual accident case report as much as possible. It is also necessary to consider that accident investigators may look for certain patterns in behaviour types and despite best practice procedures being put in place will still possibly make these types of interpretations. The cluster results similarities may be a result of these interpretations, nevertheless the HFF coding carried out in the OTS studies was as close to the original coding and all of the available information was coded, where possible within the HFF method.

The latent class clustering analysis provided detailed interpretations of the results. Despite this some of the clusters had overlap in terms of the results and caused the interpretations of the clusters to become more difficult. This was particularly the case for ‘looked but did not see’ accidents, where there were many similarities between clusters and this made it more difficult to differentiate the meaning of the results. These issues mean that the findings from the studies need to be considered alongside a detailed descriptive analysis. These issues were particularly prominent for the STATS19 data, where the cluster results were very similar in nature and did not allow a clear interpretation to be possible.

A risk analysis for individual factors was not included on the dataset. The exclusion of this information with regards to injury levels particularly reduced the interpretation possibilities of the data. Despite this the aim of the work was to find scenarios rather than identify levels of risk for individual variables. A clear risk assessment would have been beneficial to provide a more focused analysis on injury mechanisms.

As there was no possibility to have higher system level information within the analysis, this also limited the interpretations of the results. In order for a better understanding of human error, it is necessary to further understand the reasons for an error occurring. These reasons are rooted in higher level system based information such as road rules and regulations. If more detailed higher level system information could be provided a more detailed analysis could be conducted. This analysis would allow for more meaningful countermeasure implications to be discussed.

10.5.1 OTS data limitations

One of the main issues while analysing all cases in the OTS dataset was that a large number of the cases in the single and multiple vehicle accident data were missing values for the demographic variables gender and age. The large proportion in missing cases accounted for 30% of the single vehicle accident data and 35% of the multiple vehicle accident data. Each of the cases were enhanced by a cross checking procedure against STATS19 data for the specific years identifying each accident according to the accident date, time, area of accident, road user configuration and injury level. This procedure was carried out in order to decrease the number of missing cases and identify the missing demographic variables. Only cases that could be definitely identified were included in the analysis thus despite the thorough analysis of each accident case a large number of cases were not able to be identified, particularly in the all accident dataset, due to the different coding present in the two datasets. The number of cases not identified in the Pedestrian and PTW analysis chapters were less than 7% for each dataset.

A number of statistical methods were available to treat missing values, for example treating them as missing at random which is a feature in the latent class cluster analysis. Missing at random means that there is a possibility of systematic differences between observed and missing values, these values can in turn be explained by other observed variables. If the missing cases were between 5-10% then this would be a reasonable measure though the large number of missing values increased the potential for bias in the results and so only complete cases were included. A possibility of systematic bias due to accidents that did not have gender and age related variables coded also needs to be considered. An alternative is bootstrapping the data, which would carry out random sampling, estimating the missing variable's values and including them in the analysis. Bootstrapping uses the available sample data to create a large number of phantom samples and aims to replicate the population values. By using this method missing values within the measurements can be replaced based on the estimated population values. Bootstrapping relies on the available data representing the characteristics of the population adequately and provides confidence intervals for the missing values rather than point estimates of the data. The R statistics package and poLCA dataset does not allow for bootstrapping and so it was not possible to carry out this procedure.

Another possible limitation in terms of the single vehicle cluster analysis is the placement of duplicate factors in the cluster contributory factor 1 and contributory factor 2 fields. For example speed could be coded as both a contributory 1 and a contributory 2 factor for different cases, this would make it more difficult for speed to be identified as the main contributing factor 1 or factor 2 as the proportion is split between two values. The reason that this study was carried out in this manner was that there was no other way to reflect the categorical coding nature of the accident case other than including individual factors in the cluster analysis. Due to a limited number of variables and variable values being included in the cluster analysis it was not possible to enter each contributory factor value individually. If these variables would have been included a large number of the cases would not have any of the factors coded and this would have negatively influenced the cluster results. For this reason the clustering was done in the above stated manner and the descriptive analysis was used to highlight the complete coding within the analysis.

One further limitation is the timeframe for the sampling of the accident cases, as the cases are between 13-17 years old. The changing nature of the vehicle fleet and implementation of new safety measures and technologies mean that the results discussed in this study need to be further analysed with newer accident data, if available.

Due to the nature of the PTW study it was not possible to include a number of relevant factors specific to PTW accident analysis to include in the cluster analysis, such as the riders experience level and specific type of rural and urban environment. Due to the broader nature of the OTS study which was not specifically focused on PTW accidents, a number of differences to dedicated PTW studies can be highlighted. McCarthy et al. (2008) compared the OTS and MAIDS studies and found that MAIDS provided more detailed accident reconstruction data, mechanical data, human factors information and provided more detailed coding on factors specific to PTW accidents. The experience level of the rider was of particular interest in the MAIDS (2009) study, which analysed this variable in detail and if it was possible to gather this analysis level than it would have contributed to the cluster analysis significantly. Though it was not possible to use this variable, cases where the road user was inexperienced were deduced and coded for in the contributory factors section, and one of the clusters significantly emphasised inexperienced PTW riders.

Similar to the PTW study one of the main limitations of the pedestrian study was that the OTS study was not aiming to collect data specifically on pedestrian accidents and so the broad nature of the data collection process though providing a detailed case analysis did leave out some variables that would have benefitted the above study. More detailed information on pre-crash pedestrian factors could have enhanced the result findings and provided better differentiations of the clusters, or possibly a new cluster to be formed.

Although the OTS study was carried out in two specific regions in the UK, the main understanding of the results should be generalizable to the whole UK. Two separate studies were available in the literature to examine the representivity of the OTS study for all accident types and for PTW accidents alone.

A study carried out by Richards, Cookson, & Cuerden (2010) comparing the OTS study data and Great Britain national accident data between the years 2000 to 2006 found that there were differences in the two datasets with regards to road user gender and vehicle types. With regards to PTW accidents McCarthy, Walter, Hutchins, & Tong (2008) compared the first 302 OTS cases collected with the Great Britain national data. It was found that there were more severe injury cases in the OTS data but no significant differences in terms of rider age, engine size and the area type that the accident occurred in.

Both results implied that there was a significant amount of overlap between the OTS and STATS19, data nevertheless some caution needs to be exercised when considering the findings. The implications of the OTS study may also not be applicable to generalise in other country settings.

10.6 Practical applications of the study

10.6.1 Countermeasure implications

Traffic safety countermeasures aim to provide solutions to issues that road users face that lead to any type of accident. Countermeasures aim to reduce the amount of risk that a road user faces in the environment, and provide solutions that will optimally decrease the possibility of serious injury or an accident. Accident causation methods provide a description of each individual case and allow for a link to be made with possible countermeasures.

The use of accident scenarios including accident causation data enables interpretations of mass countermeasures for specific scenarios to be identified. The inclusion of relevant human factors and error data allows for a better interpretation of the combination of factors that lead to accidents. This analysis in turn allows for a better understanding of these issues.

Research studies that used macroscopic data (Depaire et al., 2008; Skyving et al., 2009) with cluster analysis methods provided limited countermeasure application discussions, rather focusing on the ability of the cluster analysis to segment accident data and group similar types of factors. This segmentation in turn was used to identify clusters of factors that caused injuries, where applicable. Previous studies that used latent class cluster analysis methods on macroscopic data (de Oña et al., 2013b; Depaire et al., 2008) identified cluster segmentation results that were mainly based on roadside physical attributes. These results required further analysis in the form of a multinomial logit model or binary networks with regards to the injury outcomes for each cluster to be used for detailed analysis findings.

In Depaire et al.’s (2008) and de Oña, López, Mujalli, & Calvo (2013b) studies the authors stated that the study was an exploratory study on whether cluster analysis could be used to segment traffic accident data, they state that cluster analysis helps with regards to interpreting data heterogeneity. The studies carried out within this thesis help underline the advantages of working with more detailed data when using LCC analysis. The selection of countermeasures based on the above stated study results would be more focused and related to specific causation factors. As the level of detail in the data increases the use of the clustering methodology provide better results and better differentiation between clusters. The inclusion of human factors related to the accident, particularly the human functional failure, and the more precise nature of the clusters in this thesis allowed for more detailed interpretations to be carried out which were not possible in the stated studies. Each cluster provided significant analysis for variable values. The significance values provided detail on the factors that occur more often in particular clusters. This allowed a differentiation of different types of accident scenarios.

An example of how to use the findings to discuss countermeasure implications is provided below.

The latent class cluster analysis of the OTS multiple vehicle accidents identified four clusters related to detection issues, as outlined in the analytic results sections. Despite the accident groupings being related to detection issues the countermeasure implications are different for the three situations. The first scenario would require a system identifying conflict possibilities based on vehicle positioning at an intersection. Education on this issue would also be beneficial. For the second scenario countermeasures related to inattention or distraction would be necessary. For the third scenario the road user's incorrect analysis of the road situation or risk taking was interpreted. The detection issues countermeasures would be based on education and enforcement rather than ADAS systems.

This separation of similar cluster groupings highlights that as the level of detail in accident data is increased, such as microscopic data compared to macroscopic data, more detailed cluster results and countermeasure implications are possible.

11 Conclusions and Recommendations

11.1 Conclusions

The aim of this thesis was to provide a procedure that would allow for accident causation data to be combined with other types of data acquired in in-depth accident analysis procedures to identify accident scenarios. The motivation for this research was to find methods that would help better understand the interactions that are present in detailed traffic accident cases.

The statistical methods and data handling procedures employed in this thesis demonstrated the possibility of causation data giving statistically relevant results when a large sample size is available. Previous work in this field concentrated on creating causation charts to analyse the data and involved analysing each case specifically for the analysis purposes (Habibovic & Davidsson, 2012; Sandin & Ljung, 2007; Sandin, 2009). This thesis provided similar detailed results using a larger number of cases with a statistical procedure.

This thesis demonstrated that when accident data has a sufficient level of detail, data mining tools can aid in the understanding of a large number of cases with regards to the causation codes present in them, though analysts need to be cautious in terms of the factors entered into the analysis so as not to alter the results. The type of factors that are included in the analysis may skew the cluster results towards one main type of factor group, thus a good balance between factors related to all possible accident factors (human, vehicle and environmental/infrastructure) are necessary to be put into the analysis. It is also possible to use macroscopic (national) databases for this process, though the limited level of detail may impair the interpretation of the results. If the data is coded in a more systematic manner then causal inferences become possible. A more detailed accident causation scheme such as the precipitating factors coding method developed for the STATS19 data would provide a better foundation for statistical analysis possibilities.

In summary, the results of the studies provided evidence in support of four main arguments:

1. Microscopic in-depth accident data is needed to understand the types of accident mechanisms that road users make in the road environment in order to provide detailed information on accident scenario segmentation.
2. A data mining approach using latent class cluster analysis can be used to develop an understanding of the different types of failure mechanisms that occur in traffic accidents.
3. Accident causation analysis of road user failures is necessary to illustrate the types of failures that road users make.
4. The development of accident scenarios helps quantify accidents and allows countermeasure indications to be made.

The main contribution to knowledge is the demonstration that the pairing of in-depth microscopic accident data with latent class clustering methods and the application of a consistent detailed description of human error allows for an analysis of clear scenarios for different traffic accident types. The advantage of this method compared to other methods of data mining is that both categorical and continuous data can be included in the same model, without the necessity of weighting the data or recoding it in such a way as to lose information within the analysis. This enables data to be included in its original coding form and for the clear relationships between factors to influence the cluster analysis method. The data analysis was initially conducted with principal component analysis, factor analysis and hierarchical cluster analysis methods but all three were not able to produce the results that were indicated in this thesis.

A further advantage and contribution to knowledge is the ability to analyse the factors relating to all vehicles involved in the each accident together rather than concentrating on the individual factors that were coded for each individual road user separately. According to the literature review this was the first type of accident analysis for multi-vehicle accidents carried out with a statistical method as a base, an advantage of which is that all interactions between road users and all factors contributing to the accident as well as the type of accident were grouped together to form clear accident patterns.

Single vehicle accident and multiple vehicle accidents were also separated analysed separately in order to clearly illustrate the differences in pre-crash factors and human error.

The advantage of the microscopic data compared to macroscopic data was also highlighted and compared in the national data analysis chapter. The resulting clusters and comparison with previous cluster analysis research results clearly demonstrated that the use of accident causation data within the cluster, paired with other types of data significantly contributed to the information that was provided by each cluster and the analysis of the accident scenarios.

In summary the main contributions that this thesis has made are:

- Demonstrating that the use of a causation model to understand human behaviour in accident situations when aiming to carry out statistical analysis is beneficial.
- Demonstrating that the use of in-depth accident data and an accident causation model allows for a detailed understanding of traffic accident scenarios to be developed when combined with a data mining tool (latent class clustering), and using a statistical method to understand differences between the clusters (chi square goodness of fit test).
- Demonstrating that in-depth accident data provides better and more detailed information than national accident data in terms of accident scenario development.

11.2 Future work

The work presented in this thesis proved the usability of accident causation data with regards to analysing accident scenarios and providing detailed data for discussion points in relation to countermeasure indications within this context. Certain limitations with regards to this analysis were present and future research could both tackle these issues as well as provide an extension of the research carried out.

The analysis carried out on the full accident dataset, powered two wheeler accident dataset and the pedestrian accident dataset helped identify certain accident scenarios. Though the accident data used was a large dataset that collected data from two areas that aimed at mirroring national data carried out in the UK and the sampling procedures carried out were within this vein, there is a possibility that the scenarios, or some aspects of them, are related to local not national issues. The limitations with regards to the data handling discussed in the general discussion if broached and tackled would allow for a better understanding of traffic accidents to be carried out. Currently two possible avenues for future research were identified in this realm.

In-depth accident research continues to be carried out in the United Kingdom. Though the UK studies are extensive in nature and detailed they would further benefit by the inclusion of one of the detailed accident causation methods (DREAM, ACASS & HFF) or a similar accident causation coding method, as it has proven difficult for the causation methods used in the OTS and STATS19 datasets to be analysed with multivariate methods for scenario development. This would help in the analysis of driver failure directly to help better detail the issues that road users face. This could be considered as an addition to the accident causation methods that have already been developed and are employed by in-depth accident research teams in the UK.

Furthermore, the application of the outlined analysis structure to naturalistic driving data could help in the data handling aspects of the data providing information related to near miss data, accident data and normal driving behaviour. This would also allow for a detection of risk factors to be possible, though in this case depending on the researcher objectives other statistical procedures may be used together with the descriptive and LCC based analysis.

A further enhancement would combine geographical information with accident causation data and the reactions of the road user collected from the vehicle would aid in further identifying failure. Such methods are mostly unavailable, since data from vehicle black boxes are seldom accessible by accident investigators. Though the level of detail used in the studies was high in terms of accident data the future possibilities of acquiring data through vehicles communicating to systems within cities before accidents and possible uses of on-board data recorders will make the data mining procedure carried out in this thesis and similar types of data mining and analysis exercises much more important in the future. In the upcoming decade the possible improvement of available new data, such as detailed driving data (acceleration, braking and steering information, driver response to stimuli, etc.) and crash data (from vehicle black-boxes), holds considerable promise for the future development of the field of accident analysis (Lord & Mannering, 2010).

Further work analysis with different statistical programs that allow latent class clustering, such as Latent Gold and SAS, could be beneficial for purposes of comparison with the results provided throughout this thesis. A comparison between the different solutions found by the different statistical analysis programs would be possible, and would allow for a better understanding of the limitations of the data if present. One of the limitations of LCC analysis is that for a relevant number of factors to be included a large number of cases must be present. The number of cases present in the datasets analysed limited the number of factors that could be entered into the analysis, though the number of factors entered for the analysis was high with a larger dataset an even more defined analysis would be possible.

The use of logistic regression models to further clarify whether the clusters identified provide new insight and information would be ideal for future considerations with regards to the handling of this data. This would be done by using accident severity analysis similar to the study using both latent class clustering and multinomial regression conducted by Depaire et al. (2008), though this study used national data from Belgium and the LCC analysis did not produce scenarios as detailed as the ones present in this thesis

Causal inferences developed from this and future data will allow for a more thorough understanding of traffic safety system issues particularly and the interaction and relationship between accident and exposure data using naturalistic methods will allow for this to occur.

References

Aarts, L., & van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention, 38*(2), 215-224. doi:10.1016/j.aap.2005.07.004

Abdel-Aty, M., & Abdelwahab, H. (2004). Analysis and prediction of traffic fatalities resulting from angle collisions including the effect of vehicles' configuration and compatibility. *Accident Analysis & Prevention, 36*(3), 457-469. doi:10.1016/S0001-4575(03)00041-1

Aczel, A. (2009). *Complete business statistics* (7th ed.) McGraw- Hill. ISBN:1935938187

af Wåhlberg, A. E. (2010). Social desirability effects in driver behavior inventories. *Journal of Safety Research, 41*(2), 99-106. doi:10.1016/j.jsr.2010.02.005

Allnutt, M. (1987). Human factors in accidents. *British Journal of Anaesthеseology, 59*(7), 856-64.

Alnaqbi, K. (2009). Investigation of pedestrian accidents analysis at signalised pedestrian crossings in Edinburgh. *Poster Presented at 4th IRTAD CONFERENCE,* 16-17 September 2009 Seoul, Korea.

Amoros, E., Martin, J. L., & Laumon, B. (2003). Comparison of road crashes incidence and severity between some French counties. *Accident Analysis & Prevention, 35*(4), 537-547. doi:10.1016/S0001-4575(02)00031-3

Atalar, D., Thomas, P., Kirk, A., Evgenikos, P., Papantoniou, P., Hermitte, T., & Van Elslande, P. (2012). *Catalogue of the current safety systems.* (No. 5.2). European DaCoTA project.

Azencott, R., Kreiss, J. P., Mougeot, M., Pastor, C., Pfeiffer, M., Siebert, S., & Zangmeister, T. (2007). *Analysis methods for accident causation studies.* (Final report No. 7.2). European TRACE project.

Badea-Romero, A., & Lenard, J. (2013). *Source of head injury for pedestrians and pedal cyclists: Striking vehicle or road? Accident Analysis &*

Prevention, 50(1), 1140–1150. doi:https://doi.org/10.1016/j.aap.2012.09.024

Baruya, A., & Finch, D. J. (1994). Investigation of traffic speeds and accidents on urban roads. *PTRC Summer Annual Meeting, 13-16 September 1994. PTRC,* London.

Bayly, M., Fildes, B., Regan, M., & Young, K. (2007). *Review of crash effectiveness of intelligent transport systems.* (Final report No. 4.1.1 – 6.2). European TRACE project.

Bédard, M., Guyatt, G. H., Stones, M. J., & Hirdes, J. P. (2002). The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. *Accident Analysis & Prevention, 34*(6), 717-727. doi:10.1016/S0001-4575(01)00072-0

Benner Jr., L. (1985). Rating accident models and investigation methodologies. *Journal of Safety Research, 16*(3), 105-126. doi:10.1016/0022-4375(85)90038-6

Berg, H., Gregersen, N. P., & Laflamme, L. (2004). Typical patterns in road-traffic accidents during driver training: An explorative Swedish national study. *Accident Analysis & Prevention, 36*(4), 603-608. doi:10.1016/S0001-4575(03)00068-X

Bjørnskau, T., & Elvik, R. (1992). Can road traffic law enforcement permanently reduce the number of accidents? *Accident Analysis & Prevention, 24*(5), 507-520. doi:10.1016/0001-4575(92)90059-R

Bjørnskau, T., Nævestad, T., & Akhtar, J. (2012). Traffic safety among motorcyclists in Norway: A study of subgroups and risk factors. *Accident Analysis & Prevention, 49*(0), 50-57. doi:10.1016/j.aap.2011.09.051

Blower, D., & Craft, R. (2005). *Large truck crash causation study analysis series.* (No. FMCSA-RI-05-035). Federal Motor Carrier Safety Administration.

Boele-Vos, M. J., Van Duijvenvoorde, K., Doumen, M. J. A., Duijvenvoorde, C. W. A. E., Louwerse, W. J. R., & Davidse, R. J. (2017). Crashes involving cyclists aged 50 and over in the Netherlands: An in-depth study. *Accident Analysis & Prevention, 105*(Supplement C), 4-10. doi:https://doi.org/10.1016/j.aap.2016.07.016

Borowsky, A., Oron-Gilad, T., Meir, A., & Parmet, Y. (2012). Drivers' perception of vulnerable road users: A hazard perception approach. *Accident Analysis & Prevention, 44*(1), 160-166. doi:10.1016/j.aap.2010.11.029

Bradbury, A. (1991). Pattern and severity of injury sustained by pedestrians in road traffic accidents with particular reference to the effect of alcohol. *Injury, 22*(2), 132-134. doi:10.1016/0020-1383(91)90074-O

Breen, J. (2012). *Managing for ambitious road safety results.* Unpublished manuscript.

Brookhuis, K. A., de Waard, D., & Jannsen, W. H. (2001). Behavioural impacts of advanced driver assistance systems-an overview. *European Journal of Transport and Infrastructure Research, 3* (Special Issue; Implementation issues of Advanced Driver Assistance Systems), 245-254.

Brorsson, B., Rydgren, H., & Ifver, J. (1993). Single-vehicle accidents in Sweden: A comparative study of risk and risk factors by age. *Journal of Safety Research, 24*(1), 55-65. doi:10.1016/0022-4375(93)90050-W

Broughton, J. (1997). A new system for recording contributory factors in road accidents. *Proceedings of the Conference Traffic Safety on Two Continents. VTI Konferens 9A, Part 5. Data Collection and Information Systems, Lisbon, Portugal,* 53-70.

Brown, I. D. (2002). *A review of the 'Look but failed to see' accident causation factor.* (Road Safety Research Report No. 60). London: Department for Transport.

Carole Millar Research. (1998). *Alcohol and the pedestrian casualty.* (Development Department Research Programme Research Findings No. 49). Edinburgh: The Scottish Office.

Carsten, O. M. J. (2007). From driver models to modelling the driver: What do we really need to know about the driver? In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments* (pp. 105-120) Springer London. doi:- 10.1007/978-1-84628-618-6_6

Carsten, O. M. J., Tight, M. R., Southwell, M. T., & Plows, B. (1989). *Urban accidents: Why do they happen?* Basingstoke, Hants: AA Foundation For Road Safety Research.

Carsten, O. M. J., & Tate, F. N. (2005). Intelligent speed adaptation: Accident savings and cost–benefit analysis. *Accident Analysis & Prevention, 37*(3), 407-416. doi:10.1016/j.aap.2004.02.007

Chandraratna, S., & Stamatiadis, N. (2009). Quasi-induced exposure method: Evaluation of not-at-fault assumption. *Accident Analysis & Prevention, 41*(2), 308-313. doi:10.1016/j.aap.2008.12.005

Chapman, P., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(2), 157-167. doi:10.1016/S1369-8478(02)00014-1

Chatfield, C., & Collins, A. J. (1980). *Introduction to multivariate analysis* (1st ed.). New York: Chapman and Hall. ISBN:0412160400

Chen, H. Y., Ivers, R. Q., Martiniuk, A. L. C., Boufous, S., Senserrick, T., Woodward, M., . . . Norton, R. (2009). Risk and type of crash among young drivers by rurality of residence: Findings from the DRIVE study. *Accident Analysis & Prevention, 41*(4), 676-682. doi:10.1016/j.aap.2009.03.005

Chorlton, K., Conner, M., & Jamson, S. (2012). Identifying the psychological determinants of risky riding: An application of an extended theory of planned behaviour. *Accident Analysis & Prevention, 49*(0), 142-153. doi:10.1016/j.aap.2011.07.003

Cicioni, G., Giuliano, G., Castellano, L., & Lattanzi, S. (1994). PC interactive computer procedures for the fast analysis of post-accident scenarios of groundwater contamination. *Advances in Engineering Software, 21*(2), 113-122. doi:10.1016/0965-9978(94)90037-X

Clabaux, N., Brenac, T., Perrin, C., Magnin, J., Canu, B., & Van Elslande, P. (2012). Motorcyclists' speed and "looked-but-failed-to-see" accidents. *Accident Analysis & Prevention, 49*(0), 73-77. doi:10.1016/j.aap.2011.07.013

Clarke, D. D., Ward, P., Bartle, B., & Truman, W. (2004). *In-depth study of motorcycle accidents.* (Road Safety Research Report No. 54). London: Department for Transport.

Clarke, D. D., Ward, P., Truman, W., & Bartle, C. (2007). *Fatal vehicle-occupant collisions: An in-depth study.* (Road Safety Research Report No. 75). London: Department for transport.

Clarke, D. D., Forsyth, R., & Wright, R. (1998). Behavioural factors in accidents at road junctions: The use of a genetic algorithm to extract descriptive rules from police case files. *Accident Analysis & Prevention, 30*(2), 223-234. doi:10.1016/S0001-4575(97)00080-8

Clarke, D. D., Ward, P., & Truman, W. (2005). Voluntary risk taking and skill deficits in young driver accidents in the UK. *Accident Analysis and Prevention, 37*, 523–529. doi:10.1016/j.aap.2005.01.007

Clarke, D. D., Ward, P., Bartle, C., & Truman, W. (2007). The role of motorcyclist and other driver behaviour in two types of serious accident in the UK. *Accident Analysis & Prevention, 39*(5), 974-981. doi:10.1016/j.aap.2007.01.002

Clifton, K. J., & Livi, A. (2005). *A smart step forward: A study of the pedestrian environment in three Maryland communities*. Princeton, New jersey: Robert Wood Johnson Foundation.

Crundall, D., Crundall, E., Clarke, D., & Shahar, A. (2012). Why do car drivers fail to give way to motorcycles at T-junctions? *Accident Analysis & Prevention, 44*(1), 88-96. doi:10.1016/j.aap.2010.10.017

Cuerden, R., Pittman, M., Dodson, E., & Hill, J. (2008). *The UK on the spot accident data collection study – phase II report*. London: Department for Transport.

Damon, N. (1958). The action program for highway safety. *The ANNALS of the American Academy of Political and Social Science, 320*(1), 15-26.

Davies, D.,G. (1999). *Research, development, and implementation of pedestrian safety facilities in the united kingdom.* (No. FHWA-RD-99-089). Georgetown Pike McLean, VA: US Department of Transportation.

Davis, G. A., & Swenson, T. (2006). Collective responsibility for freeway rear-ending accidents? An application of probabilistic causal models. *Accident Analysis & Prevention, 38*(4), 728-736. doi:10.1016/j.aap.2006.01.003

de Oña, J., López, G., & Abellán, J. (2013a). Extracting decision rules from police accident reports through decision trees. *Accident Analysis & Prevention, 50*, 1151-1160. doi:http://dx.doi.org/10.1016/j.aap.2012.09.006

de Oña, J., López, G., Mujalli, R., & Calvo, F. J. (2013b). Analysis of traffic accidents on rural highways using latent class clustering and Bayesian networks. *Accident Analysis & Prevention, 51*, 1-10. doi:10.1016/j.aap.2012.10.016

De Ridder, S., Van der Horst, R., Naing, C. L., Thomson, R., Fagerlind, H., Lanner, G., . . . Wink, W. (2006). *Roadside infrastructure for safer European roads: D04 envelope of vehicle and driver response prior to collisions*

Delhomme, P., Dobbeleer, W. d., Forward, S., & Simões, A. (2009). *Manual for designing, implementing and evaluating road safety communication campaigns.* (No. D/2009/0779/1). Brussels, Belgium: Belgian Road Safety Institute (IBSR-BIVV).

Depaire, B., Wets, G., & Vanhoof, K. (2008). Traffic accident segmentation by means of latent class clustering. *Accident Analysis & Prevention, 40*(4), 1257-1266. doi:10.1016/j.aap.2008.01.007

Department for transport. (2006). *Transport accidents and casualties*. London: The Stationery Office: Department for Transport.

Department for transport. (2007). *Transport accidents and casualties*. London: The Stationery Office: Department for Transport.

Department for transport. (2008). *Transport accidents and casualties*. London: The Stationery Office: Department for Transport.

Department for Transport. (2013). *Road casualties Great Britain 2013*. London: The Stationery Office: Department for Transport.

Department for Transport. (2010). *Motorcycling statistics: 2010*. London: The Stationery Office: Department for Transport.

di Marzo, M., Almenas, K., & Gopalnarayanan, S. (1995). Core to surge-line energy transport in a severe accident scenario. *Nuclear Engineering and Design, 154*(2), 119-132. doi:10.1016/0029-5493(94)00910-Q

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., . . . Knipling, R. R. (2006). *The 100 - car naturalistic driving study: Phase II – results of the 100 - car field experiment.* (No. DOT HS 810 593). Washington, D.C: National Highway Traffic Safety Administration.

Dozza, M., Bärgman, J., & Lee, J. D. (2013). Chunking: A procedure to improve naturalistic data analysis. *Accident Analysis & Prevention, 58*(0), 309-317. doi:10.1016/j.aap.2012.03.020

Draskóczy, M., Carsten, O., & Kulmala, R. (1998). *Road safety guidelines.* (CODE project No. B5.2). Birmingham, UK: Atkins Wootton Jeffreys.

Dunaway, K., Will, K. E., & Sabo, C. S. (2011). Chapter 17 - alcohol-impaired driving. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 231-248). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10017-7

Eby, D. W. (2011). Chapter 5 - naturalistic observational field techniques for traffic psychology research. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 61-72). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10005-0

Elvik, R., Hoye, A., Vaa, T., & Sorensen, M. (2004). *The handbook of road safety measures*. United Kingdom: Emerald Group Publishing.

Endsley, M. R. (1995). Towards a theory of situational awareness in dynamic systems. *Human Factors, 37*(1), 32-64. doi:10.1518/001872095779049543

Eshghi, A., Haughton, D., Legrand, P., Skaletsky, M., & Woolford, S. (2011). Identifying groups: A comparison of methodologies. *Journal of Data Science, 9*, 271-291.

Evans, L. (Ed.). (1991). *Traffic safety and the driver*. New York: Van Nostrand Reinhold. ISBN:0975487181

Fell, J. C. (1976). A motor vehicle accident causal system. *Human Factors, 18*, 85-94.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: Sage. ISBN:1446249185

Fildes, B., Logan, D., Fitzharris, M., Scully, J., & Burton, D. (2007). *ANCIS the first three years.* (No. 207). Victoria, Australia: Monash University Accident Research Centre.

Fillmore, K. M., Stockwell, T., Chikritzhs, T., Bostrom, A., & Kerr, W. (2007). Moderate alcohol use and reduced mortality risk: Systematic error in prospective studies and new hypotheses. *Annals of Epidemiology, 17*(5), S16-S23. doi:10.1016/j.annepidem.2007.01.005

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Michigan University: Brooks/Cole Publishing Company.

Fleury, D., & Brenac, T. (2001). Accident prototypical scenarios, a tool for road safety research and diagnostic studies. *Accident Analysis & Prevention, 33*(2), 267-276. doi:10.1016/S0001-4575(00)00041-5

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association, 97*(458), 611-631.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin. ISBN: 0313239614

Girard, Y. (1993). In safety evaluation of traffic systems. *Evaluation of Traffic Systems, Traffic Conflicts and Other Measures; Proceedings of the International Co-Operation on Theories and Concepts Traffic Safety (ICTCT) Workshop,* Salzburg: Austria.

Gitelman, V., Balasha, D., Carmel, R., Hendel, L., & Pesahov, F. (2012). Characterization of pedestrian accidents and an examination of infrastructure measures to improve pedestrian safety in Israel. *Accident Analysis & Prevention, 44*(1), 63-73. doi:10.1016/j.aap.2010.11.017

Gkikas, N. (2009). *Ergonomics of intelligent vehicle braking systems* (PhD).

Grayson, G. B., & Hakkert, A. S. (1987). Accident analysis and conflict behaviour. In J. Rothengatter, & R. de Bruin (Eds.), *Road users and traffic safety* (pp. 27-59). the Netherlands: Van Gorkum Publishers.

Greenwood, M., & Yul, G.,U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeat accidents. *Journal of the Royal Statistical Society,* (83), 256-79.

Groeger, J. A. (2002). Trafficking in cognition: Applying cognitive psychology to driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(4), 235-248. doi:10.1016/S1369-8478(03)00006-8

Groeger, J. A. (2011). Chapter 1 - how many e's in road safety? In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 3-12). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10001-3

Habibovic, A., & Davidsson, J. (2011). Requirements of a system to reduce car-to-vulnerable road user crashes in urban intersections. *Accident Analysis & Prevention, 43*(4), 1570-1580. doi:10.1016/j.aap.2011.03.019

Habibovic, A., & Davidsson, J. (2012). Causation mechanisms in car-to-vulnerable road user crashes: Implications for active safety systems. *Accident Analysis & Prevention, 49*(0), 493-500. doi:10.1016/j.aap.2012.03.022

Hagstroem, L., Fagerlind, H., Danton, R., Reed, S., Hill, J., Martensen, H., . . . Thomas, P. (2010). *Report on purpose of in-depth data and the shape of the new EU infrastructure.* (No. 2.1).European DaCoTA project.

Haque, M. M., Chin, H. C., & Huang, H. (2009). Modelling fault among motorcyclists involved in crashes. *Accident Analysis & Prevention, 41*(2), 327-335. doi:10.1016/j.aap.2008.12.010

Hautzinger, H., Pastor, C., Pfeiffer, M., & Schmidt, J. (2007). *Analysis methods for accident and injury risk studies.* (No. 7.3).European TRACE project.

Hayes, A. F. (2005). *An SPSS procedure for computing Krippendorff's alpha.* Unpublished manuscript.

Hermitte, T. (2012). *Review of accident causation models used in road accident research.* (No. 5.9).European DaCoTA project.

Herslund, M., & Jørgensen, N. O. (2003). Looked-but-failed-to-see-errors in traffic. *Accident Analysis & Prevention, 35*(6), 885-891. doi:10.1016/S0001-4575(02)00095-7

Hickford, A. J., & Hall, R. D. (2004). *Review of the contributory factors system* (Road Safety Research Report No. 43) Department for Transport: London.

Hill, J., Aldah, M., Talbot, R., Giustiniani, G., Fagerlind, H., & Jänsch, M. (2012). *Final report on the pan-European in-depth accident investigation network.* (Final report No. 2.5).European DaCoTA project.

Hill, J., & Cuerden, R. (2005). *Development and implementation of the UK on the spot accident data Collection study – phase I.* (Road Safety Research Report No. 59). London: Department for Transport.

Hole, G. (Ed.). (2007). *The psychology of driving*. London: Lawrence Erlbaum Associates. ISBN: 0805859780

Hollnagel, E. (1998). *Cognitive reliability and error analysis method*. Oxford: Elsevier Science. ISBN: 0080428487

Holloway, F. A. (1995). Low-dose alcohol effects on human behavior and performance. *Alcohol, Drugs and Driving, 11*(1), 39-56.

Howell, D. (2009). *Statistical methods for psychology* (2nd ed.). Wadsworth. CA: Cengage. ISBN: 0495597848

Huang, H., Siddiqui, C., & Abdel-Aty, M. (2011). Indexing crash worthiness and crash aggressivity by vehicle type. *Accident Analysis & Prevention, 43*(4), 1364-1370. doi:10.1016/j.aap.2011.02.010

Hummel, T. H. (1998). *Dutch pedestrian research review.* (No. R-98-7). Leidschendam: SWOV.

Hunter, W. W., Stutts, J. C., Pein, W. E., & Cox, C. L. (1995). *Pedestrian and bicycle crash types of the early 1990's.* Chapel Hill, US: UNC Highway Safety Research Centre.

Hurt, H. H., Ouellet, J. V., & Thom, D. R. (1981). *Motorcycle accident cause factors and identification of countermeasures.* (No. Volume 1). Washington DC, United States: NHTSA.

Johnston, P., Brooks, C., & Savage, H. (2008). *Fatal and serious road crashes involving motorcyclists.* (No. 20). Canberra, ACT: Department of Infrastructure, Transport, Regional Development and Local Government.

Karwat, H. (1992). *The assessment of containment codes by experiments simulating severe accident scenarios.* Luxembourg.: Office for official publications of the European Communities.

Klauer, S. G., Perez, M., & McClafferty, J. (2011). Chapter 6 - naturalistic driving studies and data coding and analysis techniques. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 73-85). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10006-2

Koustanaï, A., Boloix, E., Van Elslande, P., & Bastien, C. (2008). Statistical analysis of “looked-but-failed-to-see” accidents: Highlighting the involvement of two distinct mechanisms. *Accident Analysis & Prevention, 40*(2), 461-469. doi:10.1016/j.aap.2007.08.001

Krippendorff, K. (2004). *Content Analysis: An introduction to its methodology third edition* (2nd ed.). Thousand Oaks, CA: Sage. ISBN:1412983150

Kuge, N., Ueno, H., Ichikawa, H., & Ochiai, K. (1995). A study on the causes of rear-end collision based on an analysis of driver behavior. *JSAE Review, 16*(1), 55-60. doi:10.1016/0389-4304(94)00050-4

Kweon, Y. (2011). Chapter 8 - crash data sets and analysis. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 97-105). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10008-6

Laapotti, S., & Keskinen, E. (1998). Differences in fatal loss-of-control accidents between young male and female drivers. *Accident Analysis & Prevention, 30*(4), 435-442. doi:10.1016/S0001-4575(97)00121-8

Lai, F., Carsten, O., & Tate, F. (2012). How much benefit does intelligent speed adaptation deliver: An analysis of its potential contribution to safety and environment. *Accident Analysis & Prevention, 48*, 63-72. doi:10.1016/j.aap.2011.04.011

Lajunen, T., & Özkan, T. (2011). Chapter 4 - self-report instruments and methods. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 43-59). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10004-9

Langweider, K., Hummel, T. H., & Mueller, C. H. R. (1997). Performance of front airbags in collisions: Safety and problem areas - experience from accident research. *Proceedings of the VDI Conference on Innovative Occupant Protection and Vehicle Compatibility,* Berlin, Germany.

Lardelli-Claret, P., Luna-del-Castillo, J. d. D., Jiménez-Mejías, E., Pulido-Manzanero, J., Barrio-Anta, G., García-Martín, M., & Jiménez-Moleón, J. J. (2011). Comparison of two methods to assess the effect of age and sex on the risk of car crashes. *Accident Analysis & Prevention, 43*(4), 1555-1561. doi:10.1016/j.aap.2011.03.011

Lee, J. D. (2004). Models for transportation. In Editor-in-Chief: Charles Spielberger (Ed.), *Encyclopedia of applied psychology* (pp. 617-623). New York: Elsevier. doi:10.1016/B0-12-657410-3/00676-0

Lehto, M., & Salvendy, G. (1991). Models of accident causation and their application: Review and reappraisal. *Journal of Engineering and Technology Management, 8*(2), 173-205.

Lenard, J., & Hill, J. (2004). *Interaction of road environment, vehicle and human factors in the causation of pedestrian accidents. Proceedings of International Conference on ESAR (Expert Symposium on Accident Research), 3-4 September, Hannover, Germany*

Lenguerrand, E., Martin, J. L., & Laumon, B. (2006). Modelling the hierarchical structure of road crash data—Application to severity analysis. *Accident Analysis & Prevention, 38*(1), 43-53. doi:10.1016/j.aap.2005.06.021

Lidestam, B., Lundqvist, A., & Rönnberg, J. (2010). Concepts from research literature and practical assessment of risk awareness: The Swedish driving test from the perspective of cognitive psychology. *Transportation Research Part F: Traffic Psychology and Behaviour, 13*(6), 409-425. doi:10.1016/j.trf.2010.06.004

Lin, T. H., & Dayton, C. M. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics, 22*(3), 249-264. doi:10.3102/10769986022003249

Linzer, D. A. (2008). *poLCA: An R package for polytomous variable latent class analysis.* Unpublished manuscript.

Ljung Aust, M. (2010). Generalization of case studies in road traffic when defining pre-crash scenarios for active safety function evaluation. *Accident Analysis & Prevention, 42*(4), 1172-1183. doi:10.1016/j.aap.2010.01.006

Ljung, M. (2007). *Manual for safety net accident causation version 1.2.* (No. 5.2).European SafetyNet project.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice, 44*(5), 291-305. doi:10.1016/j.tra.2010.02.001

Lundberg, J., Rollenhagen, C., & Hollnagel, E. (2009). What-you-look-for-is-what-you-find – the consequences of underlying accident models in eight accident investigation manuals. *Safety Science, 47*(10), 1297-1311. doi:10.1016/j.ssci.2009.01.004

Mackay, G. M., Ashton, S. J., Galer, M. D. & Thomas, P. D. (1985). The Methodology of In-depth Studies of Car Crashes in Britain. *Data Collection, Methodologies and Crash Injury Reconstructions, 'Michigan*, USA. (No. 159P)

MAIDS (2009). *In-depth investigation of accidents involving powered two wheelers.* (Final Report No.2.0). Brussels: ACEM.

Malaterre, G. (1990). Error analysis and in-depth accident studies. *Ergonomics, 33*(10-11), 1403-1421.

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1-22. doi:10.1016/j.amar.2013.09.001

Maycock, G. (1995). *Contributory factors in accidents; police databases.* (No. Project Report No. PR/SE/083/95). Crowthorne: Transport Research Laboratory.

McCarthy, M.,G, Walter, L.,K, Hutchins, R., & Tong, R. (2008). *Comparative analysis of motorcycle accident data from OTS and MAIDS.* (No. PPR 168).Transport Research Laboratory.

McGwin, J., Gerald, & Brown, D. B. (1999). Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident Analysis & Prevention, 31*(3), 181-198. doi:10.1016/S0001-4575(98)00061-X

Merat, N., Jamson, H. J., Lai, F. C. H., & Carsten, O. M. J. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 54*(5), 762-771.

Michalaki, P., Quddus, M. A., Pitfield, D., & Huetson, A. (2015). Exploring the factors affecting motorway accident severity in England using the generalised ordered logistic regression model. *Journal of Safety Research, 55*(C), 89-97. doi:https://doi.org/10.1016/j.jsr.2015.09.004

Michon, J. A. (1985). A critical view of driver behavior models. what do we know, what should we do? In L. Evans, & R. Schwing (Eds.), *Human behavior and traffic safety* (pp. 485-521). New York: Plenum Press.

Michon, P. E., & Denis, M. (2001). When and why are visual landmarks used in giving directions? *Spatial Information Theory,* Berlin. 292 - 305.

Molinero, A., Evorides, H., Naing, C., Kirk, A., Tecl, J., Barrios, J. M., . . . Hermitte, T. (2008). *Accident causation and pre-accidental driving situations. part 2. in-depth accident causation analysis.* (No. 2.2). European TRACE project.

Montella, A., Aria, M., D 'Ambrosio, A., & Mauriello, F. (2012). Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. *Accident Analysis & Prevention, 49*, 58-72. doi:10.1016/j.aap.2011.04.025

Morris, A., Brace, C. L., Paulsson, R., Fagerlind, H., Ljung, M., Norin, H., . . . Hoogvelt, B. (2006). *The development of a multidisciplinary system to understand causal factors in road crashes Proceedings of the Human Factors & Ergonomics Society of Australia 42nd Annual Conference, 20-22 November,* University of Technology, UTS, Sydney.

Morris, A., Smith, M., Chambers, D., & Thomas, P. (2005). *Integrated protocols for accident research on the scene (OTS)*. Loughborough University: Loughborough

Mueller, B. A., Rivara, F., & Bergman, A. (1988). urban-rural location and the risk of dying in a pedestrian-vehicle collision. *Journal of Trauma, 28*(1), 91-4.

Naing, C., Bayer, S., Van Elslande, P., & Fouquet, F. (2007). *Which factors and situations for human functional failures? developing grids for accident causation analysis.* (No. 5.2). European TRACE project

NHTSA. (2010a). *A compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimates system.* Washington, DC: U.S. Department of Transportation.

NHTSA. (2010b). *Report to congress NHTSA's crash data collection programs.* (No. DOT HS 811 337). Washington, DC: US Department for Transport.

NHTSA. (2008). *Pedestrian crash data study*. Washington, DC: National Highway Traffic Safety Administration.

Norman, D. A. (1981). Categorization of action slips.*88* (1), 1-15. doi:10.1037/0033-295X.88.1.1

OECD. (2008). *TOWARDS ZERO ambitious road safety targets and the safe system approach.* OECD/ITF.

Olson, P.L. (2002). Driver perception - response time. *Human factors in traffic safety* (pp. 13) (Olson, P.; Dewar, L.; ed., pp. 13-20) Tucson, Arizona: Lawyers and Judges Publishing Company, Inc. ISBN:1933264241

Oppenheim, I., & Shinar, D. (2011). Chapter 15 - human factors and ergonomics. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 193-211). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10015-3

Otte, D., Jaensch, M., & Pund, B. (2007). *7 steps accident causation coding system.* The 21st International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV) June 15–18, Stuttgart, Germany.

Pai, C. (2009). Motorcyclist injury severity in angle crashes at T-junctions: Identifying significant factors and analysing what made motorists fail to yield to motorcycles. *Safety Science, 47*(8), 1097-1106. doi:10.1016/j.ssci.2008.12.007

Pai, C. (2011). Overtaking, rear-end, and door crashes involving bicycles: An empirical investigation. *Accident Analysis & Prevention, 43*(3), 1228-1235. doi:10.1016/j.aap.2011.01.004

Pai, C., & Saleh, W. (2008). Exploring motorcyclist injury severity in approach-turn collisions at T-junctions: Focusing on the effects of driver's failure to yield and junction control measures. *Accident Analysis & Prevention, 40*(2), 479-486. doi:10.1016/j.aap.2007.08.003

Peek-Asa, C., & Kraus, J. F. (1996). Injuries sustained by motorcycle riders in the approaching turn crash configuration. *Accident Analysis & Prevention, 28*(5), 561-569. doi:10.1016/0001-4575(96)00022-X

Porter, B. E. (2011). Preface. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. vii-viii). San Diego: Academic Press. doi:10.1016/B978-0-12-381984-0.10039-6

Preusser, D. F., Williams, A. F., & Ulmer, R. G. (1995). Analysis of fatal motorcycle crashes: Crash typing. *Accident Analysis & Prevention, 27*(6), 845-851. doi:10.1016/0001-4575(95)00027-5

Pund, B., Otte, D., & Jaensch, M. (2006). Systematic of analysis of human accident causation - seven steps methodology abbreviated system for the assessment of accident causation in In-depth accident studies. *ESAR-Expert Symposium on Accident Research,* 204-211.

Ranney, T. A. (1994). Models of driving behavior: A review of their evolution. *Accident Analysis & Prevention, 26*(6), 733-750. doi:10.1016/0001-4575(94)90051-5

Rasmussen, J. (1982). Human errors. A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents, 4*(2–4), 311-333. doi:10.1016/0376-6349(82)90041-4

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13*(3), 257-266.

Reason, J. (1990). *Human error*. Cambridge, UK, Cambridge University Press. ISBN:0521314194

Reason, J. (1997). *Managing the risks of organizational accidents*. Aldershot: Ashgate. ISBN:1840141050

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis & Prevention, 43*(5), 1771-1781. doi:10.1016/j.aap.2011.04.008

Rezankova, H. (2009). *Cluster analysis and categorical data.* Praha: Statistika.

Richards, D.,C. (2010). *Relationship between speed and risk of fatal injury: Pedestrians and car occupants.* (Road Safety Web Publication No.16). London: Department for Transport.

Richards, D. C., Cookson, R., & Cuerden, R. W. (2010). *Analysis of linked OTS and Stats19 records.* (No. T501S).Transport Research Laboratory.

Richardson, M., Barber, P., King, P., Hoare, E., & Cooper, D. (1997). Longitudinal driver support systems. *Proceedings of Autotech '97,* 87-97, 4-6 Nov 1997, Bury St Edmunds, UK

Riemersma, J. B. J. (1979). Perception in traffic. *Urban Ecology, 4*(2), 139-149. doi:10.1016/0304-4009(79)90015-9

Rokach, L., & Maimon, O. (2005). *Data mining with decision trees.* World Scientific Publishing Company. ISBN:981459007X

Rumar, K., & Wegman, F. (1999). *Vers le zéro accident ? L'exemple de la Suède et des Pays-Bas. Actes Du Colloque Organise a L'Occasion Du Cinquantieme Anniversaire De La Prevention Routiere, Paris.*

Sabey, B. E., & Staughton, G. C. (1975). Interacting roles of road environment, vehicle and road users in accidents. *5th International Conference of the International Association for Accident and Traffic Medicine,* London, UK.

SafetyNet. (2009). *Pedestrians & cyclists, retrieved* from http://ec.europa.eu/transport/road_safety/specialist/knowledge/pdf/pedestrians.pdf. European SafetyNet project.

Sagberg, F. (2007). *A methodological study of the driving reliability and error analysis method.* (No. 912/2007).Norwegian centre for transport research.

Salmon, P. M., Lenné, M. G., Stanton, N. A., Jenkins, D. P., & Walker, G. H. (2010). Managing error on the open road: The contribution of human error models and methods. Safety Science, 48(10), 1225-1235. doi:10.1016/j.ssci.2010.04.004

Salmon, P. M., McClure, R., & Stanton, N. A. (2012). Road transport in drift? applying contemporary systems thinking to road safety. *Safety Science, 50*(9), 1829-1838. doi:https://doi.org/10.1016/j.ssci.2012.04.011

Salmon, P. M., Read, G. J. M., & Stevens, N. J. (2016). Who is in control of road safety? A STAMP control structure analysis of the road transport system in Queensland, Australia. *Accident Analysis & Prevention,* *96*(Supplement C), 140-151. doi:https://doi.org/10.1016/j.aap.2016.05.025

Sandin, J., & Ljung, M. (2007). Understanding the causation of single vehicle crashes: A methodology for in-depth on-scene multidisciplinary case studies. *International Journal of Vehicle Science, 2*(3), 316 - 333.

Sandin, J. (2009). An analysis of common patterns in aggregated causation charts from intersection crashes. *Accident Analysis & Prevention, 41*(3), 624-632. doi:10.1016/j.aap.2009.02.015

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis & Prevention, 43*(5), 1666-1676. doi:10.1016/j.aap.2011.03.025

Scholliers, J., Bell, D., Morris, A., & Garcíad, A. B. (2014). Improving safety and mobility of vulnerable road users through ITS applications. *Transport Research Arena,* Paris.

Seiniger, P., Schröter, K., & Gail, J. (2012). Perspectives for motorcycle stability control systems. *Accident Analysis & Prevention, 44*(1), 74-81. doi:10.1016/j.aap.2010.11.018

Shankar, V., & Mannering, F. (1996). An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. *Journal of Safety Research, 27*(3), 183-194. doi:10.1016/0022-4375(96)00010-2

Shinar, D. (2007). *Traffic safety and human behavior.* Oxford: Elsevier. ISBN:0080450296

Singh, S. (2003). *Driver attributes and rear-end crash involvement propensity.* (No. DOT HS 809 540). Washington, DC: US Department for Transport

Skyving, M., Berg, H., & Laflamme, L. (2009). A pattern analysis of traffic crashes fatal to older drivers. *Accident Analysis & Prevention, 41*(2), 253-258. doi:10.1016/j.aap.2008.11.008

Staubach, M. (2009). Factors correlated with traffic accidents as a basis for evaluating advanced driver assistance systems. *Accident Analysis & Prevention, 41*(5), 1025-1033. doi:10.1016/j.aap.2009.06.014

Stutts, J., Feaganes, J., Reinfurt, D., Rodgman, E., Hamlett, C., Gish, K., & Staplin, L. (2005). Driver's exposure to distractions in their natural driving environment. *Accident Analysis & Prevention, 37*(6), 1093-1101. doi:10.1016/j.aap.2005.06.007

Summala, H. (1996). Accident risk and driver behaviour. *Safety Science, 22*(1–3), 103-117. doi:10.1016/0925-7535(96)00009-4

Sundström, G. A., & Hollnagel, E. (2006). Learning how to create resilience in .business systems. in Engineering. concepts and precepts. Ashgate: Aldershot, UK. E. Hollnagel, D. D. Woods & N. Leveson (Eds.), *Resilience engineering concepts and precepts*. Aldershot, UK: Ashgate.

Taylor, M. A. P. (2000). Network modelling of the traffic, environmental and energy effects of lower urban speed limits. *Road and Transport Research, 9*, 48-57.

Tefft, B. C. (2013). Impact speed and a pedestrian's risk of severe injury or death. *Accident Analysis & Prevention, 50*(0), 871-878. doi:10.1016/j.aap.2012.07.022

Tivesten, E., Jonsson, S., Jakobsson, L., & Norin, H. (2012). Nonresponse analysis and adjustment in a mail survey on car accidents. *Accident Analysis & Prevention, 48*(0), 401-415. doi:10.1016/j.aap.2012.02.017

Treat, J., Tumbas, N., McDonald, S., Shinar, D., Hume, R., Mayer, R., . Castellan, N. (1979). *Tri-level study of the causes of traffic accidents: Executive summary,* (No. DOT HS 805 099). Washington DC: National Highway Traffic Safety Administration.

Underwood, G., Chapman, P., Berger, Z., & Crundall, D. (2003). Driving experience, attentional focusing, and the recall of recently inspected events. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(4), 289-304. doi:10.1016/j.trf.2003.09.002

Underwood, G., Chapman, P., Bowden, K., & Crundall, D. (2002). Visual search while driving: Skill and awareness during inspection of the scene. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(2), 87-97. doi:10.1016/S1369-8478(02)00008-6

Van Elslande, P., & Fouquet, K. (2007). *Analyzing Human functional failures' in road accidents.* (No. 5.3). European TRACE project

Van Elslande, P. (2000). L'erreur humaine dans les scénarios d'accident cause ou conséquence? *Recherche - Transports - Sécurité, 66*(Supplement C), 7-31. doi:https://doi.org/10.1016/S0761-8980(00)90002-5

Wallén Warner, H., Björklund, G., Johansson, E., Ljung, A., M, & Sandin, J. (2008). *Dream manual 3.0*. Gothenburg, Sweden: Chalmers.

Wang, H. (December 4, 2014). Degrees of freedom and information criteria. Retrieved from http://www.statsblogs.com/2014/12/04/degrees-of-freedom-and-information-criteria/

Warner, H. W., & Sandin, J. (2010). The intercoder agreement when using the driving reliability and error analysis method in road traffic accident investigations. *Safety Science, 48*(5), 527-536. doi:10.1016/j.ssci.2009.12.022

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins. ISBN:0205021980

Williams, M. J., & Hoffmann, E. R. (1979). Motorcycle conspicuity and traffic accidents. *Accident Analysis & Prevention, 11*(3), 209-224. doi:10.1016/0001-4575(79)90005-8

World Health Organization. (2006). Road traffic injury prevention training manual. Geneva: World Health Organization. ISBN:9241546751

World Health Organization. (2009). *Global status report on road safety: Time for action.* Geneva: World Health Organization. ISBN: 9241563840

Wulf, G., Hancock, P. A., & Rahimi, M. (1989). Motorcycle conspicuity: An evaluation and synthesis of influential factors. *Journal of Safety Research, 20*(4), 153-176. doi:10.1016/0022-4375(89)90025-X

Yannis, G., Laiou, A., Papantoniou, P., & Gkartzonikas, C. (2013). Impact of texting on young drivers' behaviour and safety on motorways by the use of a driving simulator. *PROCEEDINGS of the 3rd International Conference on Driver Distraction and Inattention, Gothenburg, Sweden,* (No. 56-P).

Yannis, G., Golias, J., & Papadimitriou, E. (2005). Driver age and vehicle engine size effects on fault and severity in young motorcyclists accidents. *Accident Analysis & Prevention, 37*(2), 327-333. doi:10.1016/j.aap.2004.10.003

Young, M. S., Birrell, S. A., & Stanton, N. A. (2011). Safe driving in a green world: A review of driver performance benchmarks and technologies to support 'smart' driving. *Applied Ergonomics, 42*(4), 533-539. doi:10.1016/j.apergo.2010.08.012

Zegeer, C. V., & Bushell, M. (2012). Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis & Prevention, 44*(1), 3-11. doi:10.1016/j.aap.2010.12.007

Appendix A: The basic SPSS syntax to run R

The basic syntax to run the poLCA program

```
BEGIN PROGRAM R.

library(poLCA)

alldata = spssdata.GetDataFromSPSS( )

f <- cbind(Gender, Speedlimit, LightConditions, Failure, AccidentManoeuvre, Involvement,
Reaction, Contributoryfactor1, ContributoryFactor2, LAB, GenderV2, AccidentManoeuvreV2,
IV2, FV2, CF1V2, UR, V1, AgeV1, Age2V2, V2)~1

p <- poLCA(f, alldata, nclass=2, maxiter=1000, graphs=FALSE, tol=1e-10)

p2 <- poLCA(f, alldata, nclass=3, maxiter=1000, graphs=FALSE, tol=1e-10)

p3 <- poLCA(f, alldata, nclass=4, maxiter=1000, graphs=FALSE, tol=1e-10)

p4 <- poLCA(f, alldata, nclass=5, maxiter=1000, graphs=FALSE, tol=1e-10)

p5 <- poLCA(f, alldata, nclass=6, maxiter=1000, graphs=FALSE, tol=1e-10)

p6 <- poLCA(f, alldata, nclass=7, maxiter=1000, graphs=FALSE, tol=1e-10)

p7 <- poLCA(f, alldata, nclass=8, maxiter=1000, graphs=FALSE, tol=1e-10)

p8 <- poLCA(f, alldata, nclass=9, maxiter=1000, graphs=FALSE, tol=1e-10)

p9 <- poLCA(f, alldata, nclass=10, maxiter=1000, graphs=FALSE, tol=1e-10)

p10 <- poLCA(f, alldata, nclass=11, maxiter=1000, graphs=FALSE, tol=1e-10)

p11 <- poLCA(f, alldata, nclass=12, maxiter=1000, graphs=FALSE, tol=1e-10)

p12 <- poLCA(f, alldata, nclass=13, maxiter=1000, graphs=FALSE, tol=1e-10)

p13 <- poLCA(f, alldata, nclass=14, maxiter=1000, graphs=FALSE, tol=1e-10)

p14 <- poLCA(f, alldata, nclass=15, maxiter=1000, graphs=FALSE, tol=1e-10)

p15 <- poLCA(f, alldata, nclass=16, maxiter=1000, graphs=FALSE, tol=1e-10)

END PROGRAM.
```

The basic syntax to analyse the selected cluster

```
BEGIN PROGRAM R.

library(poLCA)

alldata = spssdata.GetDataFromSPSS( )
```

```
f <- cbind(Gender, Speedlimit, LightConditions, Failure, AccidentManoeuvre, Involvement,
Reaction, Contributoryfactor1, ContributoryFactor2, LAB, GenderV2, AccidentManoeuvreV2,
IV2, FV2, CF1V2, UR, V1, AgeV1, Age2V2, V2)~1

p <- poLCA(f, alldata, nclass=8, maxiter=1000, graphs=FALSE, tol=1e-10)

print(table(p$predclass))

print(p$predclass)

END PROGRAM.
```

Appendix B: Cluster analysis results and descriptive statistics for each analysis

Single vehicle cluster analysis descriptive statistics

Variable	Count	Percent
Road user gender		
Male	256	69.9
Female	110	30.1
Road user age		
0-21	69	18.9
22-29	75	20.4
30-49	155	42.3
50-65	40	11.0
66+	27	7.3
Vehicle type		
Car	292	79.8
PTW	43	11.7
Other	31	8.6
Failure mechanism		
Detection	51	13.8
Diagnosis	110	30.1
Prognosis	7	1.9
Decision	19	5.2
Execution	96	26.3
Overall	83	22.6
Area type		
Urban	166	45.3
Rural	200	54.7
Light conditions		
Day	230	62.7
Night	136	37.3
Road user contributory factor 1		
Impairment	18	4.8
Alcohol	34	9.3
Psychological factors	69	19.0
Speed	79	21.5
Breaking the law	12	3.2
Experience	18	4.9
Distraction	25	6.8
Road Condition	38	10.3

Other road factors	11	3.1
Visibility	17	4.6
Obstacle in road	14	3.9
Vehicle factors	23	6.4
No factor coded	9	2.3
Road user contributory factor 2		
Impairment	8	2.3
Psychological	91	24.8
Speed	51	13.9
Risk taking	11	2.9
Experience	23	6.4
Distraction	7	1.8
Environment	29	7.9
Other Factor	25	6.8
No factor coded	121	33.1
Emergency manoeuvre		
Brake	73	19.9
None	231	63.1
Steered	62	17.0
Speed limit		
30 mph and under	136	37.1
40-50 mph	41	11.2
60-70 mph	189	51.8
Road type		
A class	145	39.7
B class	52	14.1
Motorway	63	17.1
Minor road	106	29.0
Manoeuvre		
Going ahead	116	31.6
Left bend	94	25.6
Right Bend	67	18.4
Intersection	70	19.2
Other	19	5.2
Accident type		
Leaving lane left	163	44.4
Leaving lane right	91	24.9
Rollover	27	7.3
Collision with Obstruction/Hit parked car	26	7.1
Roundabout	34	9.4
Other	25	6.8

Single vehicle cluster analysis results

Cluster	1	2	3	4	5	6	Total
Number of cases	88	67	57	56	52	47	366
Gender							
Male	**81.5**	51.2	83.4	67.2	72.8	**87.4**	73.1
Female	18.5	**48.9**	16.6	32.8	27.3	12.6	26.9
Road user age group							
0-21	**29.5**	18.7	0,0	18.6	7.3	**47.9**	20.4
22-29	19.3	**51.7**	11,0	17.9	5.8	30.1	23.6
30-49	44.7	15.4	45.8	45.8	**70.5**	22	40.1
50-65	6.5	11.5	**21.3**	3.7	16.4	0,0	9.8
66+	0,0	2.7	**21.9**	**13.9**	0,0	0,0	6.0
Road user mode of transport							
Car	85.1	**92.2**	90.8	60.8	71.1	**97.6**	83.3
Motorcycle	7.7	7.8	3.6	**30.7**	6,0	2.4	9.7
Other	7.3	0,0	5.6	8.6	**22.9**	0,0	7.0
Road user failure mechanism							
Detection	0,0	**26.4**	0,0	**24.9**	0,0	0,0	8.9
Diagnosis	**91.5**	36.7	22.4	25.8	**91.5**	22.4	38.4
Prognosis	1.2	0,0	0,0	0,0	1.2	0,0	1.3
Decision	7.4	0,0	**33.4**	3.5	7.4	**33.4**	6.4
Execution	2.8	0,0	23.7	**28.1**	2.8	23.7	19.5
Overall	5.1	6.4	11,0	1.9	5.1	11,0	25.3
Road area type							
Urban	28.3	12.1	**54.4**	**90.5**	6.6	**82.2**	42.4
Rural	**71.7**	**87.9**	45.6	9.6	**93.4**	17.8	57.6
Light conditions							
Day	**64.7**	65.4	53.1	**81.3**	**67.5**	17.7	60.2
Night	35.3	34.6	46.9	18.8	32.5	**82.3**	39.8
Road user contributory factor 1							
Impairment	0,0	0,0	**37.9**	0,0	2.4	0,0	6.0
Alcohol	0,0	1.4	**38.2**	2.1	0,0	**40.0**	11.5

Cluster	1	2	3	4	5	6	Total
Number of cases	88	67	57	56	52	47	366
Psychological factors	**6.6**	**26.5**	19.2	**34.6**	12.2	0,0	16.7
Speed	**88.3**	15.7	0,0	19.4	3.2	22.8	30.1
Breaking the law	2.1	0,0	0,0	2.1	0,0	**23.3**	3.7
Experience	3.0	0,0	0,0	4.6	**10.3**	**13.8**	4.6
Distraction	0,0	8.1	0,0	**9.1**	7.8	0,0	4.1
Road Condition	0.0	**22.1**	0,0	15.5	**16.5**	0,0	9.1
Other road factors	0,0	11.5	0,0	1.7	1.9	0,0	2.7
Visibility	0,0	**6.6**	0,0	**9.3**	0,0	0,0	2.7
Obstacle in road	0,0	**5.8**	0,0	1.7	**11.6**	0,0	3.0
Vehicle factors	0,0	0,0	0,0	0,0	**29,0**	0,0	4.0
No factor coded	0,0	2.4	**4.8**	0,0	**5.2**	0,0	1.9
Road user contributory factor 2							
Impairment	1.2	0.0	3.6	1.7	3.9	2.1	1.9
Psychological	36.3	14.8	13.8	24.7	26.6	26.5	24.2
Speed	0,0	22.6	16.9	7.0	11.9	**48.2**	15.8
Risk taking	2.5	0,0	1.8	4.8	1.9	6.3	2.7
Experience	11.3	10.3	0,0	10.5	0,0	8.9	7.4
Distraction	0,0	1.1	3.7	**5.4**	0,0	2.2	1.9
Environment	12.1	12.7	0,0	10.3	7.4	4.4	8.6
Other Factor	7.5	11.5	5.3	8.6	4.0	1.3	6.9
No factor coded	29.1	27	**54.9**	26.8	**44.2**	0,0	30.6
Road user emergency manoeuvre							
Brake	8.6	12.2	10.2	**31.6**	**34.1**	18.1	18.0
None	72.4	48.8	74.7	62.8	61.6	62.8	63.9
Steered	18.9	**39.0**	15.2	5.6	4.3	19.1	18.2
Speed limit							
30 mph and under	16.2	11	31.6	**73.2**	4.1	**83.9**	33.2
40-50 mph	8.7	10.3	12.8	**20.6**	6.4	8.9	11.1
60-70 mph	**75.2**	**78.7**	55.7	6.2	**89.5**	7.3	55.6
Road type							
A class	37.4	38.3	50.2	38.5	41.6	36.8	40.0

Cluster	1	2	3	4	5	6	Total
Number of cases	88	67	57	56	52	47	366
B class	21.8	6.8	21.1	15.4	14.1	10.7	15.4
Motorway	13.9	23.4	9.3	0,0	**44.3**	2.2	15.6
Minor	26.8	31.5	19.4	**46.1**	0,0	**50.4**	29.0
Manoeuvre							
Going ahead	0,0	**50.0**	**39.7**	31.51	**51.3**	34.8	28.6
Left bend	**41.7**	29.6	27.8	8.04	12.3	19.1	25.0
Right Bend	**41.6**	16.1	17.8	4.34	**30.5**	10.1	15.3
Intersection	11.1	0,0	14.7	**48.8**	2.1	27.4	25.1
Other	5.5	4.3	0,0	7.3	3.9	8.6	6.0
Accident type							
Leaving lane left	**53.7**	**57.5**	48.6	29.5	46.6	39.8	47.2
Leaving lane right	28.0	29.9	**37.7**	25.9	24.6	6.4	26.4
Rollover	10.5	4.7	8.0	5.8	9.2	5.0	7.3
Collision with Obstruction/Hit parked car	0,0	1.4	2.1	8.9	9.6	**14.0**	5.1
Roundabout	2.8	0,0	2.1	**28.1**	0,0	**23.7**	8.4
Other	5.1	6.4	1.5	1.9	10.0	11.0	5.6

Multiple vehicle cluster analysis descriptive statistics

Variable	Count	Percent
Road user 1 gender		
Male	473	69.6
Female	200	30.4
Road user 1 mode of transport		
Car	479	71.1
LGV	30	4.5
HGV/BUS	42	6.3
Motorcycle	28	4.2
Pedestrian/Cycle	94	14.0
Road user 1 age group		
0-17	62	9.2
18-21	74	11.0
22-29	118	17.5
30-49	268	39.8
50-65	97	14.4
66+	31	8.0
Road user 1 failure mechanism		
Detection	318	47.4
Diagnosis	111	16.4
Prognosis	40	5.9
Decision	116	17.2
Execution	27	4.0
Overall	61	9.0
Area type		
Urban	435	64.7
Rural	238	35.3
Light conditions		
Day	524	77.9
Night	149	22.1
Road type		
A class	321	47.8
B class	110	16.3
Motorway	58	8.7
Minor	183	27.2
Road user 1 contributory factor		
Impairment	33	4.9
Alcohol	20	3.0
Psychological factors	119	17.7

Potential Risk	40	6.1
Speed	67	9.9
Breaking the law	146	21.7
Experience	22	3.3
Distraction	42	6.2
Road Condition	12	1.8
Other road factors	12	1.8
Visibility	58	8.6
Obstacle in road	4	0.6
Vehicle factors	19	2.8
None	78	11.6
Speed limit		
30 mph and under	344	51.0
40-50 mph	121	18.1
60-70 mph	208	30.9
Road user 1 manoeuvre type		
Going ahead	246	36.3
Intersection	206	30.4
Turning right	64	9.8
Turning left	21	3.1
Intersection	47	7.0
Other	89	13.3
Accident type		
Rear-end	124	18.6
Right turn against	65	9.7
Right turn same direction	32	4.8
Left turn	17	2.5
Merging road	41	6.4
Roundabout	30	4.3
Leaving lane	11	1.6
Pedestrian	95	14.1
Going into opposite lane	64	9.2
Overtaking	61	9.1
Other	132	19.7
Road user 2 gender		
Male	482	71.6
Female	191	28.4
Road user 2 age group		
0-17	25	3.7
18-21	49	7.3
22-29	127	18.9
30-49	318	47.4

50-65	124	18.4
66+	29	4.3
Road user 2 mode of transport		
Car	488	72.6
LGV	21	3.1
HGV/BUS	49	7.3
Motorcycle	66	9.8
Pedestrian/Cycle	48	7.1
Road user 2 manoeuvre		
Going ahead	357	52.9
Intersection	143	21.2
Turning	45	6.7
Overtaking	22	3.3
Slowing in traffic	63	9.4
Other	44	6.5
Road user 2 failure mechanism		
Detection	39	5.8
Diagnosis	40	6.0
Prognosis	564	83.8
Decision	16	2.4
Execution	4	0.6
Overall	3	0.4
Only Present	7	1.0
Road user 2 contributory factor		
Psychological	15	2.2
Identification	15	2.2
Risk taking	17	2.5
Traffic control	21	3.1
Atypical manoeuvres other driver	81	12.0
Illegal manoeuvres other driver	173	25.7
Other factors	20	3.0
Visibility	31	4.6
No factors coded	300	44.6

Multiple vehicle cluster analysis results

Cluster	1	2	3	4	5	6	7	8	Total
Case number	123	115	99	81	78	68	60	49	673
Road user 1 gender									
Male	61.3	66.8	69.6	**77.7**	78.1	**83.6**	73.9	56.6	69.6
Female	**38.7**	33.2	30.4	22.3	21.9	16.4	26.1	**43.4**	30.4
Road user 1 mode of transport									
Car	**90.9**	72.7	74.9	**86.7**	**85.8**	68.9	0.0	51.2	71.1
LGV	5.1	**11.4**	3.0	0.0	4.9	5.7	0.0	0.0	4.5
HGV/BUS	0.0	**12.9**	6.7	4.1	2.6	**22.4**	0.0	0.0	6.3
Motorcycle	0.0	2.9	**15.3**	**9.2**	1.2	1.5	0.0	0.0	4.2
Pedestrian/Cycle	4.0	0.0	0.0	0.0	5.5	1.5	**100.0**	**48.8**	14.0
Road user 1 age group									
0-17	0.0	0.0	6.1	6.1	3.8	0.0	**68.7**	14.2	9.2
18-21	9.3	7.9	14.2	**20.7**	12.1	0.0	5.0	**20.5**	11.0
22-29	17.0	**26.1**	23.1	16.5	16.9	12.9	1.6	16.3	17.5
30-49	**49.1**	40.1	39.6	43.8	35.4	**56.1**	11.7	28.3	39.8
50-65	17.5	19.9	11.3	8.6	18.6	19.1	5.0	8.1	14.4
66+	7.1	6.0	5.8	4.2	**13.1**	**11.8**	8.0	**12.7**	8.0
Road user 1 failure mechanism									
Detection	**65.2**	**86.8**	45.1	8.7	20.8	46.6	39.8	31.3	47.4
Diagnosis	12.4	1.0	18.0	**42.6**	**29.5**	23.6	1.7	4.1	16.4
Prognosis	0.0	6.1	**16.3**	**11.9**	0.0	3.4	3.1	6.4	5.9
Decision	19.7	0.0	**20.6**	12.0	**32.2**	5.1	18.6	**44.1**	17.2
Execution	1.6	6.2	0.0	**9.5**	0.0	**11.9**	3.4	0.0	4.0
Overall	1.1	0.0	0.0	**15.4**	**17.4**	9.4	**33.5**	14.2	9.0
Area type									
Urban	**75.0**	54.6	**82.2**	41.8	**76.2**	11.2	**84.9**	**95.9**	64.7
Rural	25.0	**45.4**	17.8	**58.2**	23.9	**88.9**	15.1	4.1	35.3
Light conditions									
Day	78.1	**87.9**	**86.1**	73.1	68.6	74.8	76.7	65.7	77.9
Night	21.9	12.1	13.9	26.9	**31.4**	25.2	23.3	**34.3**	22.1

Cluster	1	2	3	4	5	6	7	8	Total
Case number	123	115	99	81	78	68	60	49	673
Road type									
A class	34.4	**65.5**	35.7	43.3	**80.0**	40.5	26.4	56.8	47.8
B class	**23.8**	13.1	**23.5**	**26.3**	9.2	0.0	14.9	10.4	16.3
Motorway	0.0	**15.4**	0.0	0.0	0.0	**58.1**	1.7	0.0	8.7
Minor	**41.8**	6.1	**40.9**	30.3	10.8	1.3	**57.0**	32.8	27.2
Road user 1 contributory factor									
Impairment	0.0	0.0	0.0	4.7	5.1	1.6	**31.8**	10.2	4.9
Alcohol	0.0	0.0	0.0	**8.6**	**8.8**	3.0	6.7	0.0	3.0
Psychological factors	7.3	**19.5**	**37.7**	7.7	0.0	21.6	26.3	**28.7**	17.7
Potential Risk	0.0	**10.1**	8.0	1.2	**12.3**	**12.9**	1.6	2.0	6.1
Speed	2.4	9.8	8.0	**42.7**	0.0	5.6	5.0	6.1	9.9
Breaking the law	**49.9**	0.9	10.9	3.3	**58.6**	11.0	3.5	30.3	21.7
Experience	0.0	1.7	1.1	**7.3**	**7.6**	5.8	3.4	2.0	3.3
Distraction	7.6	**15.1**	6.5	6.0	0.0	0.0	3.4	4.0	6.2
Road Condition	0.0	2.9	0.0	**7.3**	0.0	4.0	0.0	0.0	1.8
Other road factors	0.0	**3.1**	**6.1**	0.0	0.0	3.7	0.0	0.0	1.8
Visibility	**20.4**	2.6	8.2	6.1	7.5	0.0	11.7	8.2	8.6
Obstacle in road	0.8	**1.7**	0.0	0.0	0.0	0.0	1.7	0.0	0.6
Vehicle factors	0.0	3.5	1.8	2.5	0.0	**16.4**	0.0	0.0	2.8
None	11.6	**28.9**	11.7	2.6	0.0	14.4	5.0	8.6	1.2
Speed limit									
30 mph and under	**74.0**	26.4	**71.8**	33.4	38.2	0.0	**81.3**	**92.2**	51.0
40-50 mph	11.7	17.7	20.1	19.8	**41.6**	10.1	13.6	7.9	18.1
60-70 mph	14.4	**55.9**	8.2	**46.8**	20.2	**89.9**	5.0	0.0	30.9
Road user 1 manoeuvre									
Going ahead	6.7	**61.1**	22.8	**82.6**	1.5	**51.7**	**48.2**	22.8	36.3
Intersection	**77.6**	25.1	6.1	5.0	**56.3**	0.0	0.0	**54.1**	30.4
Turning right	8.3	0.0	**26.2**	0.0	**36.5**	2.3	0.0	0.0	9.8
Turning left	2.3	0.9	**11.5**	4.4	2.5	0.0	0.0	0.0	3.1
Intersection	0.0	3.2	**15.5**	7.9	0.0	**24.6**	8.4	0.0	7.0
Other	4.9	9.7	17.9	0.0	3.1	**21.4**	**43.4**	**23.1**	13.3

Cluster	1	2	3	4	5	6	7	8	Total
Case number	123	115	99	81	78	68	60	49	673
Accident type									
Rear-end	0.0	**87.0**	0.0	12.6	2.4	19.0	0.0	0.0	18.6
Right turn against	**37.5**	0.0	6.3	0.0	15.2	0.0	0.0	2.1	9.7
Right turn same direction	**19.7**	0.0	6.8	0.0	1.5	0.0	0.0	0.0	4.8
Left turn	**8.1**	0.9	3.0	1.2	0.8	0.0	0.0	2.5	2.5
Merging road	0.0	0.0	**14.2**	0.0	**37.1**	0.0	0.0	0.0	6.4
Roundabout	7.0	3.1	2.1	0.0	**17.6**	0.0	0.0	2.0	4.3
Leaving lane	0.0	0.0	3.4	2.1	0.0	**8.8**	0.0	0.0	1.6
Pedestrian	0.0	0.0	0.0	0.0	0.0	0.0	**83.2**	**91.5**	14.1
Going into opposite lane	3.0	1.1	0.0	**67.8**	1.4	1.6	0.0	0.0	9.2
Overtaking	0.7	0.0	**17.7**	10.6	2.7	**41.3**	6.7	0.0	9.1
Other	24.0	7.8	**46.4**	5.7	21.3	**29.3**	10.1	2.0	19.7
Road user 2 gender									
Male	78.1	72.6	66.7	69.4	73.3	**73.9**	59.9	75.0	71.6
Female	21.9	27.4	33.3	30.6	26.7	26.1	**40.1**	25.0	28.4
Road user 2 age group									
0-17	**8.9**	0.0	0.0	2.4	0.0	0.0	1.9	**22.1**	3.7
18-21	**11.9**	2.7	5.6	3.6	12.4	7.1	6.7	8.6	7.3
22-29	23.8	12.0	25.2	10.9	23.7	27.8	8.4	16.4	18.9
30-49	38.5	55.1	47.4	48.8	45.6	50.3	56.6	36.6	47.4
50-65	13.2	23.2	19.1	**29.4**	16.0	8.8	26.4	8.3	18.4
66+	3.7	6.9	2.7	4.8	2.4	6.0	0.0	8.1	4.3
Road user 2 mode of transport									
Car	48.9	**80.4**	75.2	**84.4**	**92.5**	66.5	**90.3**	44.8	72.6
LGV	2.4	4.9	5.9	5.3	0.0	2.0	1.7	0.0	3.1
HGV/BUS	3.6	6.6	5.0	10.4	2.5	**20.1**	8.0	6.5	7.3
Motorcycle	**35.0**	6.3	6.8	0.0	1.7	11.3	0.0	0.0	9.8
Pedestrian/Cycle	10.1	1.8	7.0	0.0	3.3	0.0	0.0	**48.8**	7.1
Road user 2 manoeuvre									
Going ahead	**82.2**	12.2	41.4	**83.6**	31.5	**79.6**	**79.8**	11.5	52.9
Intersection	12.9	**28.6**	12.8	7.3	**61.0**	3.2	0.0	**52.3**	21.2

Cluster	1	2	3	4	5	6	7	8	Total
Case number	123	115	99	81	78	68	60	49	673
Turning	4.9	3.5	**23.3**	2.6	7.6	0.0	6.8	0.0	6.7
Overtaking	0.0	0.8	**12.2**	1.2	0.0	**8.8**	3.3	0.0	3.3
Slowing in traffic	0.0	**44.0**	2.3	1.2	0.0	4.9	10.1	0.0	9.4
Other	0.0	**10.9**	8.0	4.1	0.0	3.4	0.0	**36.3**	6.5
Road user 2 failure mechanism									
Detection	3.1	5.1	2.1	0.0	4.2	7.4	**20.0**	**14.4**	5.8
Diagnosis	0.0	4.4	**16.6**	0.0	0.0	**14.3**	6.7	10.1	6.0
Prognosis	**93.6**	**90.5**	69.5	**97.2**	**95.8**	74.0	70.0	61.3	83.8
Decision	2.5	0.0	**8.0**	0.0	0.0	2.9	0.0	**6.1**	2.4
Execution	0.8	0.0	0.0	0.0	0.0	1.5	1.7	2.0	0.6
Overall	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**6.1**	0.4
Only Present	0.0	0.0	**3.7**	2.9	0.0	0.0	1.7	0.0	1.0
Road user 2 contributory factor									
Psychological	0.0	0.0	**6.0**	0.0	1.4	1.5	3.4	**10.2**	2.2
Identification	2.9	2.5	2.4	0.0	1.6	2.9	0.0	6.2	2.2
Risk taking	**5.6**	0.0	**6.3**	1.2	0.0	4.4	0.0	0.0	2.5
Traffic control	2.7	1.7	**6.5**	0.0	3.1	**10.0**	0.0	0.0	3.1
Atypical manoeuvres other driver	1.0	2.4	18.1	9.7	15.2	**24.2**	**20.1**	**22.3**	12.0
Illegal manoeuvres other driver	**57.0**	0.0	10.4	22.6	**71.4**	6.0	6.8	20.6	25.7
Other factors	0.0	**8.6**	0.0	0.0	1.3	4.7	8.4	2.0	3.0
Visibility	0.0	0.9	3.1	0.0	0.0	0.0	**25.1**	**11.8**	4.6
No factors coded	27.5	**84.0**	47.1	**65.2**	4.7	46.3	**36.3**	26.9	44.6

STATS19 cluster analysis descriptive statistics

Variable	Count	Percent
Road type		
Roundabout	4289	7.7
One way street	814	1.5
Dual carriageway	8825	15.9
Single Carriageway	40675	73.3
Slip road	584	1.1
Unknown	291	0.5
Speed limit		
30 mph	31363	56.5
40 mph	5531	10.0
50 mph	1560	2.8
60 mph	12167	21.9
70 mph	4857	8.8
Junction detail		
Roundabout	5183	9.3
Mini roundabout	516	0.9
T or staggered junction	18470	33.3
No junction	18259	32.9
Slip road	937	1.7
Crossroads	6435	11.6
Four or more arms	974	1.8
Private drive/entrance	2399	4.3
Other junction	2305	4.2
Junction control		
Authorised person	102	0.2
Traffic signal	5542	10.0
Stop sign	564	1.0
Give way	31024	55.9
Uncontrolled	18247	32.9
Road user 1 mode of transport		
Cycle	1662	3.0
PTW	3771	6.8
Car	44103	79.5
LGV	896	1.6
HGV	5047	9.1
Road user 2 mode of transport		
Cycle	3500	6.3
PTW	4668	8.4
Car	41771	75.3
LGV	1171	2.1
HGV	4365	7.9
Light conditions		
Day	51546	92.9
Night	3933	7.1
Road user 1 manoeuvre		
Turning left	2141	3.9
Turning right	11058	19.9
Waiting	2377	4.3
Lane change	1780	3.2
Overtaking	2815	5.1
Going ahead left bend	3146	5.7
Going ahead right bend	1912	3.4
Going ahead	22410	40.4

Other	7841	14.1
Road user 2 manoeuvre		
Turning left	1044	1.9
Turning right	4099	7.4
Waiting	6348	11.4
Lane change	583	1.1
Overtaking	1823	3.3
Going ahead left bend	1747	3.1
Going ahead right bend	2993	5.4
Going ahead	29327	52.9
Other	7518	13.6
Road user 1 gender		
Male	39665	71.5
Female	15815	28.5
Road user 2 gender		
Male	37764	68.1
Female	17716	31.9
Road user 1 age group		
0-17	2823	5.1
18-21	7621	13.7
22-29	10644	19.2
30-49	21422	38.6
50-65	8777	15.8
66+	4193	7.6
Road user 2 age group		
0-17	706	1.3
18-21	5246	9.5
22-29	9603	17.3
30-49	26119	47.1
50-65	10793	19.5
66+	3013	5.4
Road user 1 contributory factor 1		
Road environment	5502	9.9
Vehicle defects	483	0.9
Injudicious action	8037	14.5
Error or reaction	24944	45.0
Impairment/Distraction	2645	4.8
Behaviour/Inexperience	3565	6.4
Vision affected by external	2390	4.3
Pedestrian only	1438	2.6
Special codes	799	1.4
No factor coded	5675	10.2
Road user 1 contributory factor 2		
Road environment	1773	3.2
Vehicle defects	216	0.4
Injudicious action	3967	7.2
Error or reaction	16593	29.9
Impairment/Distraction	1365	2.5
Behaviour/Inexperience	3984	7.2
Vision affected by external	2085	3.8
Pedestrian only	1062	1.9
Special codes	382	0.7
No factor coded	24053	43.4
Road user 2 contributory factor 1		
Road environment	1845	3.3
Vehicle defects	192	0.3
Injudicious action	2672	4.8

Error or reaction	7829	14.1
Impairment/Distraction	715	1.3
Behaviour/Inexperience	1512	2.7
Vision affected by external	1150	2.1
Pedestrian only	609	1.1
Special codes	371	0.7
No factor coded	38587	69.6
Road user 2 contributory factor 2		
Road environment	353	0.6
Vehicle defects	37	0.1
Injudicious action	920	1.7
Error or reaction	3370	6.1
Impairment/Distraction	290	0.5
Behaviour/Inexperience	768	1.4
Vision affected by external	508	0.9
Pedestrian only	277	0.5
Special codes	86	0.2
No factor coded	48871	88.1
Road type		
Motorway	4289	7.7
A(M)	814	1.5
A	8825	15.9
B	40675	73.3
C	584	1.1
Unclassified	291	0.5

STATS19 cluster analysis results

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
Road type														
Roundabout	0.1	0.1	0.1	7.2	0.1	**92.4**	0.1	0.2	0.0	0.2	0.1	0.2	8.3	7.7
One way street	**2.2**	**2.0**	1.2	**2.0**	0.2	0.0	**2.6**	0.2	0.4	**2.4**	3.0	1.3	1.2	1.5
Dual carriageway	2.8	2.1	5.8	9.9	12.5	2.8	**36.6**	97.2	1.7	3.7	8.9	9.5	**69.6**	15.9
Single Carriageway	**94.4**	**94.8**	**91.9**	**79.9**	**86.9**	4.6	59.8	0.0	**97.0**	**92.6**	**87.2**	**87.6**	0.0	73.3
Slip road	0.1	0.5	0.3	0.5	0.3	0.2	0.2	2.0	0.1	0.5	0.2	0.3	**20.8**	1.1
Unknown	0.5	0.5	**0.8**	0.6	0.1	0.0	**0.8**	0.4	**0.8**	0.5	**0.7**	**1.1**	0.2	0.5
Speed limit														
30 mph	**94.5**	**62.1**	50.3	**70.2**	26.2	**57.3**	**76.6**	1.2	20.6	**84.4**	65.0	53.4	7.9	56.5
40 mph	2.9	8.1	10.1	**13.4**	**19.7**	**16.3**	**17.1**	5.3	5.9	5.0	9.2	**11.5**	**11.5**	10.0
50 mph	0.0	2.2	**3.0**	2.4	**5.7**	**3.9**	3.0	**5.6**	2.2	0.9	2.1	2.9	**8.0**	2.8
60 mph	2.6	**27.6**	**36.6**	13.6	**46.3**	15.9	2.6	1.3	**71.2**	9.7	**23.8**	**31.0**	8.2	21.9
70 mph	0.0	0.0	0.0	0.5	2.1	6.7	0.7	**86.7**	0.1	0.0	0.0	1.2	**64.5**	8.8
Junction detail														
Roundabout	0.4	3.9	0.0	**9.6**	0.2	**91.3**	0.2	0.0	0.0	1.4	0.0	0.0	**37.0**	9.3
Mini roundabout	0.4	0.8	0.0	**1.2**	0.0	**8.5**	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.9

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
T or staggered junction	**69.0**	**58.2**	0.0	**53.3**	**68.8**	0.0	25.5	0.0	0.0	**65.5**	0.0	0.0	10.9	33.3
No junction	0.0	0.1	**100.0**	0.0	0.0	0.0	0.0	**100.0**	**100.0**	0.0	**100.0**	**100.0**	0.0	32.9
Slip road	0.2	1.3	0.0	1.2	1.7	0.0	0.3	0.0	0.0	0.6	0.0	0.0	**39.1**	1.7
Crossroads	**15.7**	**16.3**	0.0	**16.3**	10.0	0.0	**59.8**	0.0	0.0	11.2	0.0	0.0	1.5	11.6
Four or more arms	1.4	1.6	0.0	**3.6**	1.4	0.1	**11.2**	0.0	0.0	1.2	0.0	0.0	1.5	1.8
Private drive/entrance	**7.4**	**7.9**	0.0	**6.8**	**11.8**	0.0	0.1	0.0	0.0	**11.5**	0.0	0.0	2.4	4.3
Other junction	**5.6**	**10.0**	0.0	**8.0**	**6.1**	0.0	2.9	0.0	0.0	**7.6**	0.0	0.0	**7.5**	4.2
Junction control														
Authorised person	0.2	**0.3**	0.0	0.3	0.2	**0.3**	0.2	0.0	0.0	0.3	0.1	0.0	**0.7**	0.2
Traffic signal	4.5	2.9	0.0	**16.1**	1.0	7.3	**89.6**	0.0	0.0	4.0	0.0	0.0	9.8	10.0
Stop sign	**1.5**	**1.6**	0.0	**1.3**	**2.5**	0.4	**2.4**	0.0	0.0	1.3	0.0	0.0	0.5	1.0
Give way	**93.8**	**95.2**	0.0	**82.3**	**96.4**	**92.1**	7.8	0.0	0.0	**94.4**	0.0	0.1	**89.0**	55.9
Uncontrolled	0.0	0.0	**100.0**	0.0	0.0	0.0	0.0	**100.0**	**100.0**	0.0	**99.9**	**100.0**	0.0	32.9
Road user 1 mode of transport														
Cycle	0.2	0.0	**5.7**	2.3	0.6	2.0	2.8	0.2	0.7	**26.6**	2.7	1.8	0.7	3.0
PTW	0.2	3.6	**10.9**	**8.1**	0.6	5.6	3.1	3.6	**9.3**	**43.7**	**3.8**	5.9	2.7	6.8
Car	**93.3**	**85.2**	75.4	**80.8**	**89.5**	**80.6**	**85.1**	66.6	**80.1**	27.6	**80.7**	**80.0**	77.9	79.5

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
LGV	1.3	**1.8**	1.2	**2.7**	1.1	1.6	**2.7**	0.5	1.3	0.5	3.1	**3.1**	1.4	1.6
HGV	5.0	9.4	6.9	6.0	8.3	**10.2**	6.4	**29.1**	8.6	1.6	9.8	9.2	**17.2**	9.1
Road user 2 mode of transport														
Cycle	**18.0**	0.8	1.0	**10.9**	2.1	**13.3**	2.8	0.9	1.2	0.9	16.2	**10.0**	1.7	6.3
PTW	**21.1**	1.1	0.9	8.3	**17.1**	**12.1**	6.2	2.6	2.1	1.1	20.4	**12.2**	3.9	8.4
Car	55.8	**88.8**	**86.0**	72.8	72.5	67.5	**82.9**	**76.0**	**80.3**	**89.8**	**51.3**	69.8	**81.1**	75.3
LGV	**2.8**	**2.2**	2.0	1.5	1.2	1.7	**3.0**	0.9	**2.8**	1.9	**4.7**	1.4	0.7	2.1
HGV	2.2	7.2	**10.2**	6.5	7.2	5.5	5.1	**19.6**	**13.7**	6.3	**7.5**	6.7	**12.6**	**7.9**
Light conditions														
Day	**93.7**	91.8	92.5	**93.2**	**93.4**	**93.2**	93.5	92.3	89.0	**97.2**	94.7	91.2	93.3	92.9
Night	6.3	**8.2**	7.5	6.9	6.6	6.9	6.5	7.7	**11.0**	2.8	**5.3**	**8.9**	6.7	7.1
Road user 1 manoeuvre														
Turning left	**10.9**	3.4	0.4	3.1	**6.1**	**7.1**	2.0	0.1	0.1	3.9	1.4	0.7	3.1	3.9
Turning right	**57.5**	3.4	0.7	11.8	**75.4**	9.3	**34.6**	0.1	0.1	9.2	10.3	2.4	3.6	19.9
Waiting	3.9	3.4	0.9	**16.1**	2.9	**5.3**	4.4	1.3	0.0	0.5	4.4	**10.4**	**7.7**	4.3
Lane change	0.6	0.3	0.8	0.3	1.0	**4.4**	1.6	**24.0**	0.1	0.9	5.0	1.3	**15.5**	3.2
Overtaking	1.0	4.4	**9.0**	3.2	0.3	0.9	1.0	**7.3**	4.4	**22.6**	12.4	**5.8**	2.4	5.1

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
Going ahead left bend	0.0	**8.4**	2.1	1.4	0.0	1.6	0.1	1.6	**60.6**	1.2	**0.5**	3.8	2.7	5.7
Going ahead right bend	0.1	**3.8**	1.9	2.6	0.1	1.9	0.3	1.0	**30.4**	1.0	**0.5**	**8.7**	0.9	3.4
Going ahead	10.6	**58.3**	**72.6**	**50.4**	4.1	**47.1**	**45.4**	**51.9**	3.4	**56.5**	26.9	**42.7**	**42.7**	40.4
Other	**15.4**	14.7	11.6	11.2	10.2	**22.3**	10.6	12.8	1.0	4.2	38.7	**24.4**	**21.4**	14.1
Road user 2 manoeuvre														
Turning left	1.1	**2.6**	0.7	**5.1**	0.3	**4.2**	1.8	0.0	0.1	**5.7**	0.2	0.8	1.8	1.9
Turning right	1.7	**8.5**	2.1	**25.9**	0.9	**9.9**	**11.3**	0.1	0.2	**33.0**	0.8	3.1	1.9	7.4
Waiting	2.2	**31.2**	**14.6**	4.9	0.8	**17.0**	**12.9**	9.2	1.6	10.6	6.8	3.0	**22.4**	11.4
Lane change	0.2	0.2	0.4	**2.0**	0.2	**1.4**	0.7	**5.2**	0.0	0.6	0.3	**2.3**	**4.4**	1.1
Overtaking	**5.7**	0.3	0.9	**5.1**	**4.7**	1.4	0.6	**4.8**	0.3	1.7	13.8	**8.1**	2.5	3.3
Going ahead left bend	0.8	3.0	0.2	2.0	2.4	1.1	0.4	1.3	**26.4**	0.8	**1.1**	**10.0**	1.5	3.1
Going ahead right bend	2.5	**6.6**	0.2	1.2	4.4	3.1	0.4	0.9	**50.9**	1.1	**1.3**	4.4	1.2	5.4
Going ahead	**81.1**	30.8	51.4	40.9	**85.9**	48.4	**60.8**	**59.8**	17.6	34.3	59.0	51.0	45.7	52.9
Other	4.7	**16.8**	**29.7**	13.0	0.4	13.5	11.2	**18.7**	2.9	12.4	16.9	**17.4**	**18.7**	13.6
Road user 1 gender														
Male	62.2	71.4	**76.4**	66.9	63.4	69.6	70.7	**78.2**	**77.0**	**92.1**	**72.4**	69.5	**73.8**	71.5
Female	**37.8**	28.6	23.6	**33.1**	**36.6**	**30.4**	29.4	21.8	23.0	7.9	27.6	**30.5**	26.2	28.5

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
Road user 2 gender														
Male	**72.6**	59.9	63.9	**70.9**	**71.4**	65.5	68.1	**72.4**	69.0	64.3	77.3	**73.0**	67.0	68.1
Female	27.4	27.4	**40.1**	**36.1**	29.2	28.6	34.5	31.9	27.6	31.0	35.8	22.7	27.0	33.0
Road user 1 age group														
0-17	1.0	2.0	**10.2**	1.7	1.7	3.1	2.9	0.9	**5.9**	**38.7**	**2.9**	1.8	1.6	5.1
18-21	12.0	**18.3**	**18.1**	8.0	10.5	11.0	**16.3**	10.0	**19.1**	**18.1**	**9.5**	8.6	10.5	13.7
22-29	19.6	**21.2**	19.7	16.9	14.3	18.2	**22.4**	**21.0**	**22.9**	15.7	**16.3**	16.2	**21.8**	19.2
30-49	**39.9**	37.6	34.8	**46.4**	34.5	**40.5**	38.4	**46.9**	36.8	22.5	40.7	**44.5**	**42.1**	38.6
50-65	**17.9**	13.0	11.4	**21.1**	**20.5**	**18.7**	14.4	**17.3**	11.5	4.1	21.2	**22.7**	**17.5**	15.8
66+	**9.6**	8.1	6.0	5.9	**18.5**	**8.6**	5.5	3.9	3.9	1.1	9.4	6.2	6.4	7.6
Road user 2 age group														
0-17	**3.1**	0.0	0.1	**4.5**	0.8	0.9	0.4	0.1	0.2	0.1	2.3	**4.9**	0.1	1.3
18-21	**12.5**	6.7	7.1	**15.1**	**10.0**	9.3	8.0	6.0	7.0	9.3	10.2	**17.4**	6.9	9.5
22-29	**20.1**	14.9	14.9	**20.7**	**17.4**	17.5	**20.3**	17.2	12.1	**18.0**	16.6	**18.9**	16.5	17.3
30-49	44.5	**51.3**	**49.8**	37.9	45.8	48.3	46.9	**51.0**	**48.2**	46.3	**49.5**	39.8	**50.0**	47.1
50-65	15.2	**21.8**	**22.1**	14.5	**20.5**	18.9	19.2	**22.0**	**25.9**	**20.5**	**17.1**	13.1	**21.6**	19.5
66+	4.6	5.3	**6.1**	**7.2**	**5.7**	5.1	5.2	3.7	**6.8**	**5.9**	**4.4**	6.0	4.9	5.4

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
Road user 1 contributory factor 1														
Road environment	3.9	**16.9**	**15.5**	2.1	5.7	6.8	6.4	8.3	**37.7**	4.3	**1.3**	7.7	6.3	9.9
Vehicle defects	0.2	**1.1**	**1.6**	0.1	0.3	0.7	0.6	**1.6**	**1.2**	**2.5**	**0.2**	0.0	**1.5**	0.9
Injudicious action	9.7	**19.1**	**18.2**	3.6	11.3	**14.9**	**27.8**	13.8	**15.7**	**23.3**	**4.2**	5.5	14.4	14.5
Error or reaction	**66.2**	38.4	35.8	10.4	**71.2**	**59.1**	43.9	**45.8**	24.7	43.6	64.0	14.9	**53.0**	45.0
Impairment/Distraction	2.0	**8.3**	**12.0**	0.2	1.1	3.4	4.8	**7.3**	5.0	3.1	1.4	0.5	4.4	4.8
Behaviour/Inexperience	4.0	**7.8**	**11.0**	1.1	4.3	5.9	6.0	5.8	**9.7**	**13.9**	**3.8**	1.7	6.4	6.4
Vision affected by external	**8.2**	3.5	3.2	2.0	**5.2**	3.2	2.4	**6.1**	3.3	2.8	7.4	3.4	3.2	4.3
Pedestrian only	**4.0**	**2.8**	0.6	0.1	0.7	**2.9**	**4.8**	1.5	0.9	**3.7**	12.5	0.1	1.3	2.6
Special codes	1.3	**1.5**	**1.6**	0.6	0.3	0.7	**2.2**	**2.2**	1.4	1.6	4.0	0.6	1.5	1.4
No factor coded	0.6	0.7	0.6	**79.9**	0.1	2.4	1.3	7.7	0.4	1.2	1.3	**65.6**	8.0	10.2
Road user 1 contributory factor 2														
Road environment	1.0	**5.6**	**5.2**	0.3	1.5	1.4	1.5	2.0	**16.3**	1.1	**0.4**	1.3	2.0	3.2
Vehicle defects	0.1	**0.6**	**0.7**	0.0	0.1	0.3	0.4	**0.6**	**0.9**	**0.9**	**0.0**	0.0	0.3	0.4
Injudicious action	2.2	**13.3**	**11.4**	0.7	2.5	6.7	**8.5**	**8.3**	**14.7**	**9.4**	**0.4**	1.9	7.0	7.2
Error or reaction	**40.5**	29.3	27.4	1.3	**52.4**	**35.1**	**31.9**	28.2	25.0	**30.2**	32.2	7.3	30.6	29.9
Impairment/Distraction	1.3	**3.8**	**5.4**	0.1	1.4	1.9	2.4	**3.8**	2.4	2.4	1.8	0.1	1.5	2.5

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
Behaviour/Inexperience	5.4	**7.9**	**11.1**	0.0	5.8	6.7	**9.5**	5.5	**9.9**	**15.8**	**4.6**	1.8	7.8	7.2
Vision affected by external	**7.9**	3.1	2.6	0.2	**5.7**	3.1	2.0	3.8	**4.9**	1.9	**5.2**	2.9	2.1	3.8
Pedestrian only	**3.2**	**2.0**	0.2	0.0	0.4	1.7	**3.9**	1.2	0.7	**2.9**	9.6	0.1	0.8	1.9
Special codes	0.5	0.7	**0.9**	0.0	0.4	0.4	**1.1**	**1.1**	0.7	**1.0**	2.5	0.1	0.2	0.7
No factor coded	38.0	33.7	35.2	**97.5**	29.9	42.7	38.9	**45.7**	24.5	34.3	43.3	**84.7**	**47.6**	43.4
Road user 2 contributory factor 1														
Road environment	2.3	2.1	1.7	**7.1**	3.3	1.4	1.9	2.1	**9.3**	0.8	**0.4**	**17.0**	1.8	3.3
Vehicle defects	0.4	0.1	0.1	**1.1**	0.2	0.5	0.1	0.4	0.2	0.4	0.2	**1.4**	0.5	0.3
Injudicious action	4.9	1.1	0.9	**18.9**	**5.2**	2.8	**6.5**	3.6	2.6	1.0	2.7	**16.2**	3.2	4.8
Error or reaction	8.1	7.2	6.3	**59.0**	6.4	**14.1**	7.9	12.4	4.1	**15.6**	7.1	**44.7**	**15.3**	14.1
Impairment/Distraction	**1.6**	0.6	0.7	**4.1**	0.6	**1.6**	0.3	1.4	0.3	0.6	0.8	**5.4**	0.6	1.3
Behaviour/Inexperience	2.3	2.0	2.4	**5.4**	1.9	2.4	2.1	2.1	3.0	2.8	1.1	**8.4**	3.3	2.7
Vision affected by external	2.1	1.9	1.5	**2.8**	1.6	0.8	0.8	2.1	**4.3**	**3.6**	**1.6**	**4.9**	1.0	2.1
Pedestrian only	**1.9**	0.7	0.5	0.4	0.1	1.2	**2.0**	0.7	0.5	**2.2**	3.8	0.3	0.8	1.1
Special codes	0.6	0.6	0.8	**1.1**	0.3	0.6	0.7	0.8	0.3	0.5	0.8	**1.6**	0.5	0.7
No factor coded	**76.0**	**83.8**	**85.1**	0.2	**80.4**	**74.7**	**77.8**	**74.4**	**75.4**	**72.5**	**81.6**	0.0	**73.2**	69.6
Road user 2 contributory factor 2														

	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
	7911	7700	6701	4460	4316	4094	4044	3839	3295	3007	2419	1997	1698	55474
Road environment	0.2	0.3	0.1	**1.9**	0.3	0.3	0.2	0.4	**2.7**	0.0	**0.0**	**4.5**	0.0	0.6
Vehicle defects	0.1	0.0	0.0	**0.3**	0.0	0.1	0.0	0.1	0.0	0.0	0.1	**0.4**	0.1	0.1
Injudicious action	0.9	0.2	0.1	**9.5**	0.7	0.6	0.8	1.2	1.1	0.2	0.0	**10.2**	1.4	1.7
Error or reaction	1.5	0.8	0.4	**40.3**	1.5	3.6	2.7	3.9	1.3	4.1	0.1	**32.0**	4.8	6.1
Impairment/Distraction	0.4	0.0	0.0	**3.1**	0.0	0.3	0.0	0.5	0.2	0.0	0.0	**3.7**	0.1	0.5
Behaviour/Inexperience	0.9	0.2	0.1	**7.3**	0.4	0.7	0.6	0.8	0.2	0.8	0.0	**9.7**	1.4	1.4
Vision affected by external	0.4	0.4	0.0	**3.6**	0.8	0.3	0.2	0.4	**1.6**	1.1	**0.0**	**6.1**	0.4	0.9
Pedestrian only	**0.9**	0.2	0.0	0.3	0.0	0.5	**1.0**	0.3	0.2	**1.1**	2.0	0.4	0.3	0.5
Special codes	0.0	0.1	0.0	**0.7**	0.0	0.2	0.0	0.1	0.1	0.1	0.0	**1.5**	0.1	0.2
No factor coded	**94.6**	**97.7**	**99.4**	33.1	**96.2**	**93.6**	**94.4**	**92.3**	**92.7**	**92.5**	**97.8**	31.5	**91.5**	88.1
Road type														
Motorway	0.1	0.1	0.1	7.2	0.1	**92.4**	0.1	0.2	0.0	0.2	0.1	0.2	8.3	7.7
A(M)	**2.2**	**2.0**	1.2	**2.0**	0.2	0.0	**2.6**	0.2	0.4	**2.4**	3.0	1.3	1.2	1.5
A	2.8	2.1	5.8	9.9	12.5	2.8	**36.6**	**97.2**	1.7	3.7	8.9	9.5	**69.6**	15.9
B	**94.4**	**94.8**	**91.9**	**79.9**	**86.9**	4.6	59.8	0.0	**97.0**	**92.6**	**87.2**	**87.6**	0.0	73.3
C	0.1	0.5	0.3	0.5	0.3	0.2	0.2	**2.0**	0.1	0.5	0.2	0.3	**20.8**	1.1
Unclassified	0.5	0.5	**0.8**	0.6	0.1	0.0	**0.8**	0.4	**0.8**	0.5	**0.7**	**1.1**	0.2	0.5

PTW cluster analysis descriptive statistics

Variable	Count	Percent
Gender		
Male	384	89.7
Female	44	10.3
Rider failure mechanism		
Detection	68	15.9
Diagnosis	61	14.4
Prognosis	206	48.1
Decision	50	11.7
Execution	19	4.6
Overall	23	5.4
Area type		
Urban	267	62.5
Rural	161	37.5
Light conditions		
Day	339	79.2
Night	89	20.8
Rider contributory factor		
Physical/physiological	89	20.9
Risk taking	90	20.9
Experience	13	3.0
Distraction	11	2.6
Road Condition	12	2.7
Traffic Condition	87	20.3
Visibility Impaired	14	3.2
Other Environmental factors	6	1.4
Vehicular factor	7	1.6
No Factor	100	23.4
Other road user emergency failure mechanism		
Yes	155	36.2
No	273	63.8
Rider level of involvement		
Primary Contributory	213	49.8
Secondary Contributory	33	7.7
Not Contributory	182	42.5
Road type		
A class	209	48.9
B class	66	15.4
Motorway	23	5.3
Minor	130	30.3

Rider age range		
0-18	72	16.9
19-25	88	20.6
26-45	192	44.9
46-65	59	13.8
66+	16	3.8
Speed limit		
30 mph and under	194	45.4
FortyFifty mph	125	29.2
SixtySeventy mph	108	25.3
PTW engine size		
50	71	16.5
51-250	84	19.6
250+	274	63.9
Interacting variable		
Detection	174	40.7
Prognosis	96	22.5
Decision	39	9.1
Single Vehicle	95	22.2
Other	24	5.6
Accident type		
Leaving lane	107	25.0
Rear-end	41	9.6
Changing lane	39	9.2
Overtaking	36	8.5
Right turn	114	26.6
Left turn	15	3.5
Intersection	31	7.2
Other	45	10.5

PTW cluster analysis results

Cluster	1	2	3	4	5	6	7	Total
Number of cases	122	77	75	45	42	36	31	428
Rider gender								
Male	83.0	90.9	92.6	91.1	90.5	95.5	96.6	89.7
Female	**17.0**	9.1	7.4	8.9	9.6	4.5	3.4	10.3
Rider failure mechanism								
Detection	2.5	2.9	1.3	**74.8**	**53.8**	4.1	13	15.9
Diagnosis	0.9	**50.8**	2.6	**25.3**	13.8	0.0	6.8	14.4
Prognosis	**96.6**	0.0	**96.1**	0.0	0.0	33.2	12.9	48.1
Decision	0.0	**25.1**	0.0	0.0	14.6	**62.8**	6.8	11.7
Execution	0.0	0	0.0	0.0	4.6	0.0	**56.6**	4.6
Overall	0.0	**21.2**	0.0	0.0	**13.2**	0.0	3.9	5.4
Area type								
Urban	**87.8**	42.1	26	61.8	**90**	**88.5**	35.2	62.5
Rural	12.2	**57.9**	**74**	38.2	10	11.5	**64.8**	37.5
Light conditions								
Day	74.2	83.5	79.4	85.1	71.8	90.1	77.2	79.2
Night	25.8	16.5	20.6	14.9	28.2	9.9	22.9	20.8
Rider contributory factor								
Physical/physiological	7.9	**33.5**	15.6	30.6	**35.7**	30.5	8.0	20.9
Risk taking	1.5	**57.8**	4.1	25.5	23.0	**49.9**	3.6	20.9
Inexperience	0.7	2.6	0.0	2.0	16.7	3.3	3.2	3.0
Distraction	0.8	1.3	0.0	**12.8**	3.0	0.0	6.5	2.6
Road Condition	0.8	0.0	0.0	0.0	0.0	0.0	**34.5**	2.7
Traffic Condition	**40.5**	4.9	**34**	9.0	7.7	3.0	0	20.3
Visibility Impaired	**4.9**	0.0	0.0	2.2	**9.3**	5.8	3.1	3.2
Other Environmental factors	0.0	0.0	0.0	0.0	0.0	0.0	18.8	1.4
Vehicular factor	0.0	0.0	0.0	0.0	4.7	0.0	**15.7**	1.6
No Factor	**43**	0.0	**46.4**	18.0	0.0	7.5	6.5	23.4
Road user emergency manoeuvre								

Cluster	1	2	3	4	5	6	7	Total
Number of cases	122	77	75	45	42	36	31	428
Yes	32.6	30.2	**46.5**	35.3	37.9	34.9	41.4	36.2
No	67.5	69.8	53.6	64.7	62.1	65.1	58.6	63.8
Level of involvement								
Primary Contributory	0.0	**100**	6	**100**	**93.1**	62.7	**81**	49.8
Secondary Contributory	7.1	0.0	10.9	0.0	4.6	**34.1**	6.3	7.7
Not Contributory	**92.9**	0.0	**83.1**	0.0	2.4	3.2	12.8	42.5
Road type								
A class	37.4	44.3	**67.9**	62.7	24.1	**67.6**	51.9	48.9
B class	**22.9**	10.2	12.3	17.3	17.1	10.9	6.6	15.4
Motorway	0.0	1.2	**13.5**	4.3	0.0	0.0	**31.8**	5.3
Minor	**39.8**	**44.3**	6.3	15.7	**58.8**	21.5	9.7	30.3
Rider age group								
0-18	18.6	0.0	3.7	2.9	**88.0**	15.9	9.5	16.9
19-25	21.8	**34.8**	20.6	15.3	9.5	21.1	3.2	20.6
26-45	41.4	46.0	**51.6**	53.4	2.5	52	**76.3**	44.9
46-65	13.9	16.9	15.9	**28.5**	0.0	3.1	10.9	13.8
66+	4.3	2.4	**8.2**	0.0	0.0	7.8	0.0	3.8
Speed limit								
30 mph and under	**76.0**	40.7	0.0	32.7	**84**	48.2	9.8	45.4
40-50 mph	24.0	13.1	**50.2**	33.4	16.1	**48.5**	28.9	29.2
60-70 mph	0.0	**46.2**	**49.8**	34.0	0.0	3.4	**61.2**	25.3
PTW engine size								
50	**24.5**	4.8	6.0	0.0	**66.6**	7.2	6.5	16.5
51-250	23.1	19.5	15.8	15.8	28.4	16.7	12.6	19.6
250+	52.5	**75.8**	**78.2**	**84.2**	5.1	76.1	81.0	63.9
Other vehicle failure type								
Detection	**76.4**	0.0	**79.5**	11.1	2.4	42.3	0.0	40.7
Prognosis	0.0	25.2	1.3	**80.4**	**64.1**	23.8	13.2	22.5
Decision	**14.0**	1.3	10.9	0.0	6.8	**27.5**	0.0	9.1
Single Vehicle	0.0	**73.4**	0.0	8.5	24.3	3.6	**74.2**	22.2

Cluster	1	2	3	4	5	6	7	Total
Number of cases	122	77	75	45	42	36	31	428
Other	**9.7**	0.0	8.2	0.0	2.4	2.8	12.6	5.6
Accident type								
Leaving lane	3.3	**83.7**	5.2	0.0	25.2	0.0	**77.2**	25.0
Rear-end	4.2	0.0	7.9	**41.8**	**22.0**	0.0	6.3	9.6
Changing lane	2.1	0.0	**31.4**	8.9	6.9	14.5	3.2	9.2
Overtaking	3.0	5.8	9.7	**29**	9.4	10.6	0.0	8.5
Right turn	**53.6**	0.0	22.6	16	15.5	**49.2**	0.0	26.6
Left turn	**6.4**	0.0	7.4	0.0	0.0	4.7	0.0	3.5
Intersection	**12.9**	0.0	11.0	0.0	13.6	3.5	0.0	7.2
Other	14.5	10.5	4.9	4.3	7.4	17.6	13.3	10.5

Pedestrian cluster analysis descriptive statistics

Variable	Count	Percent
Road user gender		
Male	184	74.2
Female	63	25.8
Road user failure mechanism		
Detection	64	26.2
Diagnosis	15	6.1
Prognosis	134	54.1
Decision	15	6.3
Execution	5	2.2
Overall	12	5.1
Area type		
Urban	220	89.1
Rural	28	10.9
Light conditions		
Day	172	69.7
Night	75	30.3
Pedestrian contributing factor		
Alcohol/Impairment	35	14.5
Young age/Pedestrian playing	53	21.4
Psychological state	51	20.6
Risk taking	27	10.9
Other driver	22	9.1
Visibility impaired	20	8.4
None	37	15.2
Manoeuvre		
Going Ahead	142	57.4
Traffic lights	43	17.6
Intersection	9	3.8
Overtaking	17	6.9
Pedestrian Crossing	11	4.8
Other	23	9.6
Road user emergency manoeuvre		
Yes	134	54.1
No	113	45.9
Road user level of involvement		
Primary Contributory	70	28.5
Secondary Contributory	14	5.8
Not Contributory	162	65.7
Road user age group	51	20.7

0-18		
19-25	85	34.5
26-45	63	25.8
45-65	9	3.8
66+	37	15.3
Pedestrian age group		
0-12	67	27.1
13-17	39	15.9
19-29	40	16.5
30-65	66	26.7
66+	34	13.9
Speed limit		
30 mph and under	218	88.2
Over 30 mph	29	11.8
Pedestrian failure mechanism		
Detection	55	22.4
Diagnosis	9	3.8
Prognosis	53	21.6
Decision	44	17.8
Execution	3	1.4
Overall	81	33.1
Accident type		
Crossing road	85	34.6
Crossing intersection	47	19.2
Crossing between cars	52	21.1
Vehicle crash	22	9.0
Other	40	16.1
Pedestrian gender		
Male	154	62.4
Female	93	37.7

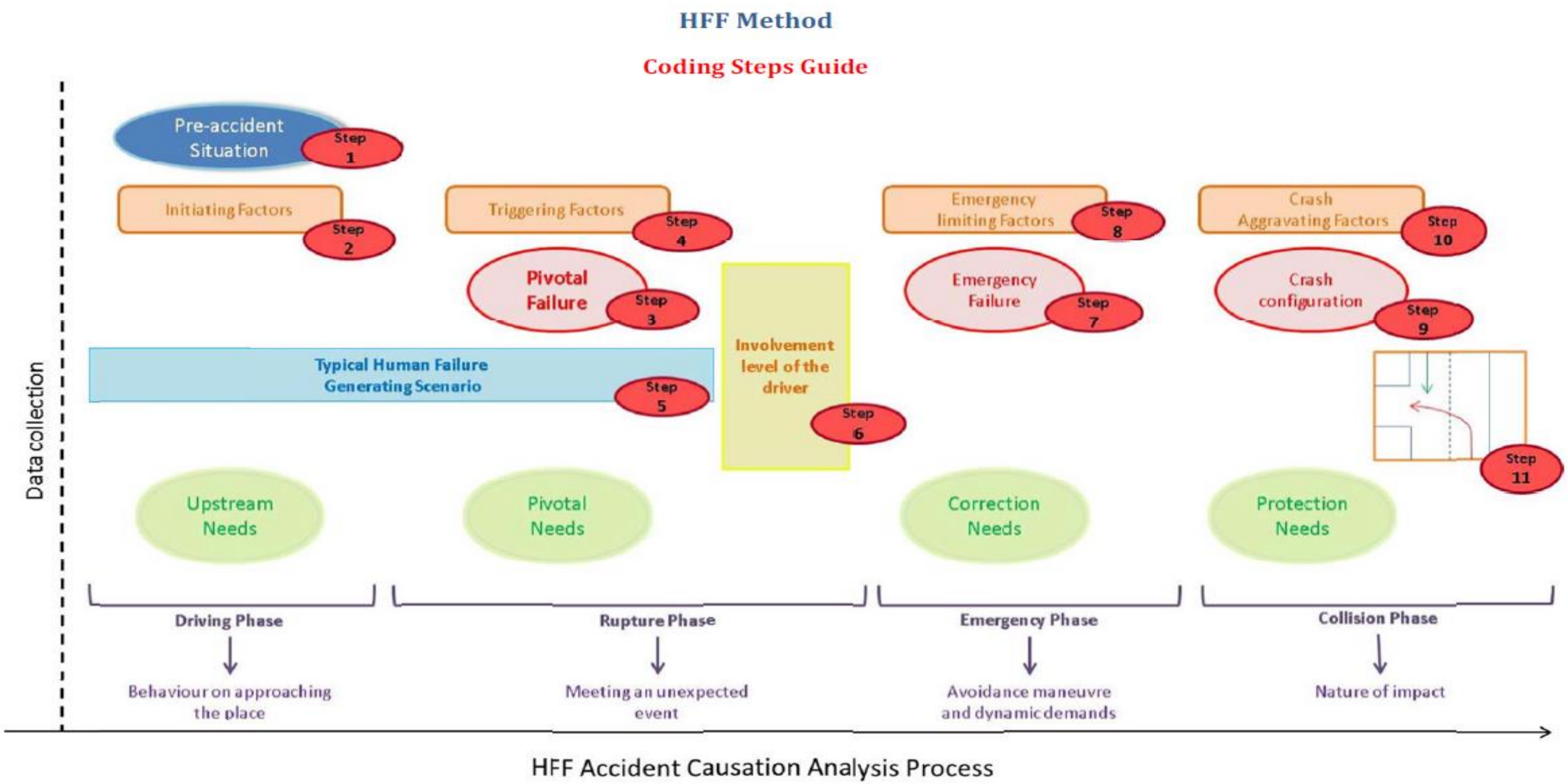
Pedestrian cluster analysis results

Cluster	1	2	3	4	Total
Number of cases	78	60	60	50	248
Road user gender					
Male	**74.8**	73.5	82.2	66.5	74.2
Female	25.2	26.5	17.8	33.5	25.8
Road user failure mechanism					
Detection	18.6	**46.4**	16.9	22.9	26.2
Diagnosis	8.3	8.3	5.9	2.0	6.1
Prognosis	**73.1**	5.4	**68.9**	**69.2**	54.1
Decision	0.0	**16.6**	6.7	2.0	6.3
Execution	0.0	**6.6**	0.0	2.0	2.2
Overall	0.0	**16.6**	1.7	2.0	5.1
Area type					
Urban	92.1	93.4	86.6	84.4	89.1
Rural	8.0	6.6	13.4	15.6	10.9
Light conditions					
Day	**83.1**	71.8	42.3	**81.4**	69.7
Night	16.9	28.2	**57.7**	18.6	30.3
Pedestrian contributory factor					
Alcohol/Impairment	1.6	0.0	**56.4**	0.0	14.5
Young age/Pedestrian playing	8.5	0.0	0.0	**77.1**	21.4
Psychological state	**59.2**	9.9	2.6	10.5	20.6
Risk taking	10.6	0.0	**33.0**	0.0	10.9
Other driver	0.0	**36.5**	0.0	0.0	9.1
Visibility impaired	**16.5**	3.4	3.5	10.1	8.4
None	3.7	**50.2**	4.5	2.3	15.2
Road user manoeuvre					
Going Ahead	**59.3**	30.4	57.1	**82.8**	57.4
Traffic lights	**20.3**	18.2	**32.1**	0.0	17.6
Intersection	**5.2**	6.6	3.3	0.0	3.8
Overtaking	5.8	6.6	0.0	**15.1**	6.9

Cluster	1	2	3	4	Total
Number of cases	78	60	60	50	248
Pedestrian Crossing	2.7	**13.3**	3.2	0.0	4.8
Other	6.8	**24.9**	4.4	2.1	9.6
Road user emergency manoeuvre					
Yes	49.3	32.0	59.6	**75.5**	54.1
No	50.7	**68.0**	40.4	24.5	45.9
Road user level of involvement					
Primary Contributory	0.0	**100.0**	7.9	6.0	28.5
Secondary Contributory	4.6	0.0	**10.9**	7.8	5.8
Not Contributory	**95.4**	0.0	**81.2**	**86.3**	65.7
Road user age group					
0-18	17.1	22.0	26.2	17.5	20.7
19-25	19.0	29.8	**47.2**	42.0	34.5
26-45	**37.7**	18.3	10.9	36.3	25.8
45-65	6.4	**6.6**	0.0	2.0	3.8
66+	19.8	**23.2**	15.8	2.2	15.3
Pedestrian age group					
0-12	23.1	11.7	3.3	**70.3**	27.1
13-17	21.5	9.9	2.4	**29.7**	15.9
19-29	**24.7**	13.3	**27.9**	0.0	16.5
30-65	21.2	**41.5**	**44.2**	0.0	26.7
66+	9.5	**23.7**	**22.3**	0.0	13.9
Speed limit					
30 mph and under	86.8	91.7	80.4	94.0	88.2
Over 30 mph	13.2	8.3	**19.6**	6.0	11.8
Pedestrian failure mechanism					
Detection	**66.2**	2.2	5.1	16.2	22.4
Diagnosis	5.1	**10.0**	0.0	0.0	3.8
Prognosis	0.0	**82.9**	3.4	0.0	21.6
Decision	22.1	3.3	**39.4**	6.4	17.8
Execution	0.0	1.7	0.0	**4.0**	1.4
Overall	6.6	0.0	**52.2**	**73.5**	33.1

Cluster	1	2	3	4	Total
Number of cases	78	60	60	50	248
Pedestrian behaviour					
Crossing road	38.6	13.7	**47.8**	38.2	34.6
Crossing intersection	**27.9**	18.2	**30.5**	0.0	19.2
Crossing between cars	22.2	3.4	3.4	**55.4**	21.1
Vehicle crash	1.3	**31.5**	3.3	0.0	9.0
Other	10.1	**33.2**	15.0	6.3	16.1
Pedestrian gender					
Male	52.5	46.9	**79.9**	**70.2**	62.4
Female	**47.6**	**53.1**	20.2	29.8	37.7

Appendix C: HFF and LAB coding sheets



HFF
Method

Sheet N° 1: Pre-accident situation

Coding Step: 1

A. Stabilised	
A.1 Going ahead	**A.1.1** Going ahead on a straight road
	A.1.2 Going ahead on a left bend
	A.1.3 Going ahead on a right bend
B. Intersection	
B.1 On approach	**B.1.1** Approaching a 'give way' intersection
	B.1.2 Approaching a 'stop' intersection
	B.1.3 Approaching a 'traffic signal' intersection
	B.1.4 Approaching intersection where road user has right
B.2 Stopped	**B.2.1** Stopped at a 'give way' intersection
	B.2.2 Stopped at a 'stop' intersection
	B.2.3 Stopped at a 'traffic signal' intersection
	B.2.4 Stopped in road/ turning lane waiting to turn
B.3 Going ahead	**B.3.1** Going straight on at a 'give-way' intersection
	B.3.2 Going straight on at a 'stop' intersection
	B.3.3 Going straight on at a 'traffic signal' intersection
	B.3.4 Crossing intersection where road user has right of
	B.3.5 Travelling on roundabout (not turning on/off)
	B.3.6 Travelling on slip-road (not turning on/off)
B.4 Turning	**B.4.1** Turning across traffic at a 'give-way' intersection
	B.4.2 Turning across traffic at a 'stop' intersection
	B.4.3 Turning across traffic at a 'traffic signal'
	B.4.4 Turning across traffic from main road into side road
	B.4.5 Turning away from traffic at a 'give-way'
	B.4.6 Turning away from traffic at a 'stop' intersection
	B.4.7 Turning away from traffic at a 'traffic signal'
	B.4.8 Turning away from traffic from main road into side
C.	
C.1 Overtaking	**C.1.1** Overtaking stationary vehicle on left
	C.1.2 Overtaking stationary vehicle on right
	C.1.3 Overtaking moving vehicle on left
	C.1.4 Overtaking moving vehicle on right

C.2 Changing lane	**C.2.1** Moved into lane on left (NOT overtaking)
	C.2.2 Moved into lane on right (NOT overtaking)
C.3 Slowing	**C.3.1** Stopping (not at junction)
	C.3.2 Parking (roadside)
C.4 Starting	**C.4.1** Starting (not at junction)
	C.4.2 Leaving parking space (roadside)
C.5 Turning (not at intersection)	**C.5.1** Turning across traffic from main road into private
	C.5.2 Turning away from traffic from main road into
	C.5.3 Turning across traffic out of private drive
	C.5.4 Turning away from traffic out of private drive
C.6 Reversing	**C.6.1** Reversing
C.7 U-turn	**C.7.1** U-turn
C.8 In wrong direction	**C.8.1** Driving in wrong direction (e.g. down a one-way
D.	
D.1 Parked	**C.1.1** Parked
D.2 Stopped in traffic queue	**D.2.1** Stopped in traffic queue
D.3 Pedestrian crossing	**D.3.1** Approaching pedestrian crossing
	D.3.1 Stopped at pedestrian crossing
D.4 Railway crossing	**D.4.1** Approaching railway crossing
	D.4.2 Stopped at railway crossing

HFF Method

Sheet N° 2: Factors

Coding Step: 2-4 and 8

User related Factors	A. User State	*1. Physical/ Physiological*	***A.1.1*** *Medical condition*	*Heart condition/Epilepsy/Other brain condition/Respiratory condition/Blood condition/Other*
			A.1.2 Pre-existing impairment	Hearing/Visual/Physical disability/Other impairment
			A.1.3 Behavioural	Linked to age
		2. Psycho-physiological condition	**A.2.1** Substances taken - alcohol	Above 'legal' limit/ Below 'legal'
			A.2.2 Substances taken - drugs	Illegal drugs
			A.2.3 Substances taken -	Correctly used medication/ Misused
			A.2.4 Emotional	Upset/Angry/Anxious/Happy/Other emotion
			A.2.5 Fatigue	Physical/Mental
			A.2.6 In a hurry	In a hurry
			A.2.7 Panic	The road user is overwhelmed by the situation
		3. Internal conditioning of performed task	**A.3.1** Right of way status	Rigid attachment to the right of way status
			A.3.2 Excessive confidence	Excessive confidence in signs given to others
			A.3.3 Identification of	Identification of potential risk about only part of the situation

		A.3.4 Overall time constraint	Affected to the journey
		A.3.5 Situational time constraint	Affected to a maneuver
		A.3.6 Trivialization of the	Neglect the potential risk associated with the situation, notably for well-known and usual situations
		A.3.7 Illusion of visibility	The road user is confident in the fact that he has been seen by the other (often the case for less conspicuous road users: PTW riders,
		A.3.8 Lights off	By night or during the day for vehicles which must put them on (PTW, cars for countries where it is compulsory)
	4. Risk taking	**A.4.1** Illegal Speed	Illegal/Erratic/Other
		A.4.2 Legal Speed but inappropriate	Legal but inappropriate to situation constrains
		A.4.3 Vehicle positioning	In front/Lateral/Other
		A.4.4 Traffic control	Signs disobeyed/Signals disobeyed /Markings disobeyed/Othe
		A.4.5 'Eccentric' motives	Testing a vehicle/Thrill-seeking/Competing/'Stunt'/Unspecified eccentric motives
		A.4.6 Atypical acceleratio	Levels of acceleration which can surprise the other road users (specifically for motorbikes)
		A.4.7 Atypical overtaking	Overtaking on the wrong lane / filtering / gymkhana
		A.4.8 Excess of caution	Too much caution affected to the driving activity
B. Experience	1. Little/None	**B.1.1** Driving	Learner/New driver/Infrequent driver/Other
		B.1.2 Route	New route/Road type/New road/Road feature/Driving on the left/Driving on the right/Other

		B.1.3 Vehicle	New vehicle/ Transmission type/
C. Attention			Right hand drive vehicle/ Other vehicle
		B.1.4 Environment	Night driving/City driving/Country driving/Driving in snow/Driving in fog/Driving in wet or flood/Driving in ice/Other
		B.1.5 Driving	Change in driving rules/Other
	2. Over - experienced	**B.2.1** Route	Route in general/Road type/New road/Road feature/Other
		B.2.2 Vehicle	New vehicle/ Transmission type/Other vehicle feature
		B.2.3 Environment	Night driving/City driving/Country driving/Driving in snow/Driving in fog/Driving in wet or flood/Driving in ice/other
	1. Attention disturbances	**C.1.1** Distraction outside vehicle*	Police/Animal in road/ Sunlight or sunset/ People in roadway/ Crash scene/Other perceived danger/Road construction/ Searching for directional
		C.1.2 Distraction within vehicle*	Adjusting radio/ Adjusting cassette/ Adjusting CD/ Other occupant/ Moving object in vehicle/ sing or viewing device integral to vehicle/ Using other device brought into vehicle/ Adjusting climate controls/ Eating/Drinking/ Cell
		C.1.3 Distraction within user*	Lost in thought/ Medical

Environment related Factors			
	D. Road Condition	**D. 1** Contaminants: Wet/Flood/Snow	Wet/Flood/Snow
		D.2 Contaminants: Ice/Frost	Ice/Frost
		D. 3 Contaminants: Oil/Diesel	Oil/Diesel
		D. 4 Contaminants: Sand/Gravel/Mud	Sand/Gravel/Mud
		D. 5 Surface defects	Potholes/Cracks/Bumps
		D. 6 Surface type	Asphalt/Concrete/Untreated/Cobbles /Brick/Other
	E. Road Geometry	**E. 1** Bend(s)	Left/Right/Wide/Tight/Multiple bends
		E. 2 Slope(s)	Decline/Incline/Multiple slopes
		E. 3 Road width	Wide/Narrow/Single lane/Multiple lanes/Change in width
		E. 4 Adverse camber	Left/Right
		E. 5 Traffic calming	Road hump/Speed table/Throttle/Chicane
		E. 6 Temporary road layout	Roadworks/Other
		E. 7 Misleading/complex	Misleading/Complex
		E. 8 Speed-inciting layout	Bend in road/Straight road/Gradient/Wide road/Continuity effect
		E.9 Monotonous Layout	Ex: Motorway
	F. Traffic Condition	**F.1** Difficulties of obtaining an insertion slot	Traffic dense, erratic, at high speed
		F.2 Other road user(s) : Absence of clues	Absence of clues to manoeuvre
		F.3 Other road user(s) : Ambiguity of clues to manoeuvre	Ambiguity of clues to manoeuvre
		F.4 Other road user(s) : Atypical manoeuvres	Atypical manoeuvres
		F.5 Illegal road user(s) manoeuvres	No respect of Traffic light/ stop / signal

		F.6 Disruptive behaviour of another	Low speed/ hesitant behaviour
		F.7 Being drawn into manoeuvre	Passenger/Vehicle ahead/Vehicle behind/Pedestrian/Cyclist
	G. Visibility Impaired	**G.1** Road lighting	Type/Colour/Intensity/No lighting
		G.2 Vehicle lighting	Type/Colour/Beam type/No lighting
		G.3 Day/night	Daylight/Darkness/Dusk/Dawn
		G.4 Sun glare	Direct from sun/Reflection from wet road
		G.5 Weather	Rain/Fog or mist/Snow/Hail
		G.6 Smoke	Vehicle/Nearby fire/Other
		G.7 Terrain profile	Bend/Slope/Side slope(s)/Other
		G.8 Other vehicle(s)	High vehicle/Wide vehicle/Parked vehicle/Vehicle stopped in traffic/Other
		G.9 Roadside objects	Overhanging tree(s)/ Overhanging shrubbery/Sign(s)/Bridge structures/Barrier(s)/Wall(s)/Boundary fence(s)/Other
		G.10 Glare	Vehicle lights of other user/ sun
	H. Traffic Guidance	**H.1** Traffic signs/signals - Insufficient	Signs present but insufficient/Signals present but insufficient/Signs
		H.2 Traffic signs/signals – Maintenance	Signs damaged/Signals damaged/Signs poorly maintained/Signals poorly maintained/Signs positioned incorrectly/Signals positioned incorrectly/Other
		H.3 Traffic signs/signals – Unexpected	Signs replaced/Signals replaced/Signs new/Signals new/Other
		H.4 Traffic signs/signals – Inappropriate	Signs inappropriate/Signals inappropriate/Signs confusing/Signals confusing /Other

		H.5 Road markings (visual/tactile) - Insufficient	Visual markings present but insufficient/Tactile markings present but insufficient/Visual markings absent/Tactile markings absent

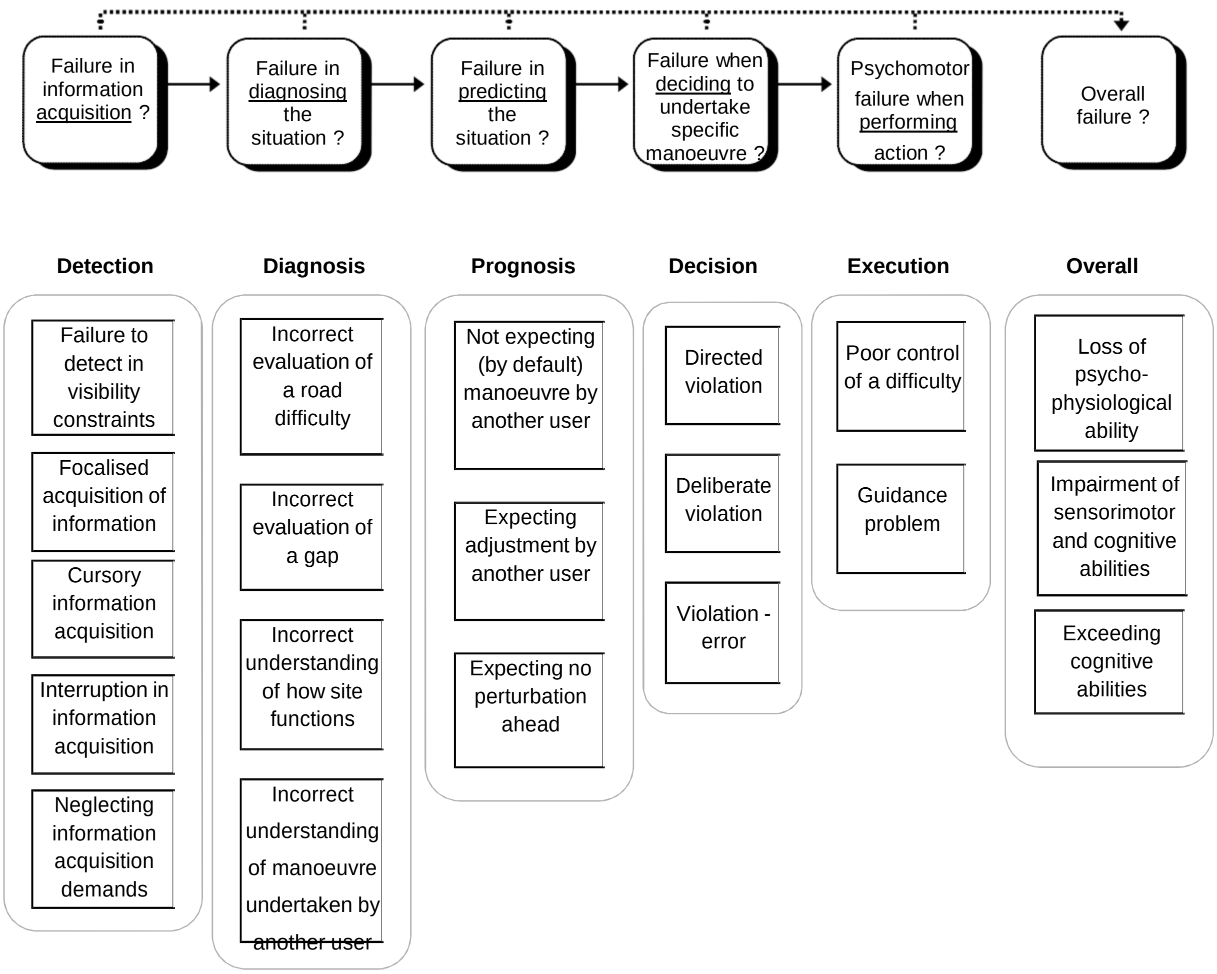
vehicle related Factors			
	J. Electro-mechanical	**J.1** Steering	Partial failure/Total failure
		J.2 Brakes	Partial failure/Total failure
		J.3 Engine	Partial failure/Total failure
		J.4 Suspension	Partial failure/Total failure
		J.5 Electrical/electronics	Partial failure/Total failure
	K. Maintenance	**K.1** Windscreen/Glass	Front chipped/ Front cracked/ Front misted/Front dirty / Front scratched/Rear chipped/ Rear cracked/Rear misted/ Rear dirty/Rear
		K.2 Tyre(s)	Incorrect type/Air pressure/ Tread/ Blow-
		K.3 Exterior lights	Headlight type/Headlight bulb needs replacing/Headlight cracked/Headlight broken cover/ Rear light type/ Rear light bulb needs replacing/ Rear light cracked/ Rear light broken cover/ Brake light type/ Brake light bulb needs replacing/ Brake light cracked/ Brake light broken cover/ Indicator type/ Indicator bulb needs
		K.4 Interior lights	Fuel light/Oil light/Water light/Parking brake light/Other dashboard

	L. Design	**L.1** Visibility	A-pillar(s)/B-pillar(s)/C-pillar(s)/Steering wheel blocking view/Rear view
		L.2 Auditory	Auditory warnings confusing
		L.3 Displays	Colour/Size/Confusing information
		L.4 Controls	Colour/Size/Confusing information/Reach
	M. Load	**M.1** Heavy	On vehicle/Within vehicle/Other
		M.2 Uneven	On vehicle/Within vehicle/Other
		M.3 Visibility obstructed	On vehicle/Within vehicle/Other

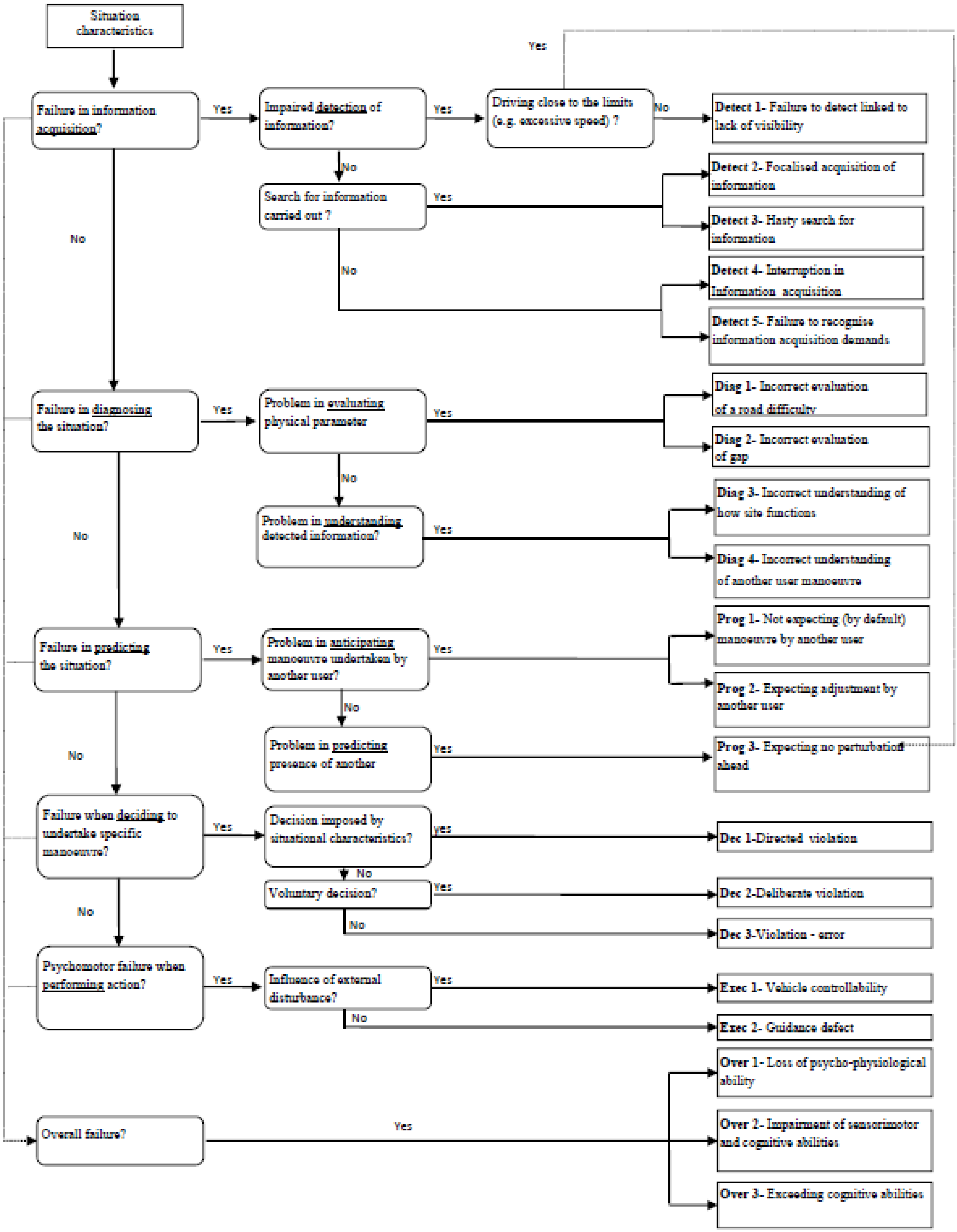
HFF Method

Sheet N° 3: Pivotal Functional Failure

Coding Step: 3



Delineation of functional failures found in In-depth accident data



Situation characteristics
Failure in information acquisition?
Yes
Impaired detection of information?
Yes
Driving close to the limits (e.g. excessive speed) ?
No
Detect 1- Failure to detect linked to lack of visibility
Yes
No
Search for information carried out ?
Yes
Detect 2- Focalised acquisition of information
Detect 3- Hasty search for information
No
Detect 4- Interruption in Information acquisition
Detect 5- Failure to recognise information acquisition demands
No
Failure in diagnosing the situation?
Yes
Problem in evaluating physical parameter
Yes
Diag 1- Incorrect evaluation of a road difficulty
Diag 2- Incorrect evaluation of gap
No
Problem in understanding detected information?
Yes
Diag 3- Incorrect understanding of how site functions
Diag 4- Incorrect understanding of another user manoeuvre
No
Failure in predicting the situation?
Yes
Problem in anticipating manoeuvre undertaken by another user?
Yes
Prog 1- Not expecting (by default) manoeuvre by another user
Prog 2- Expecting adjustment by another user
No
Problem in predicting presence of another
Yes
Prog 3- Expecting no perturbation ahead
No
Failure when deciding to undertake specific manoeuvre?
Yes
Decision imposed by situational characteristics?
yes
Dec 1-Directed violation
No
Voluntary decision?
Yes
Dec 2-Deliberate violation
No
Dec 3-Violation - error
No
Psychomotor failure when performing action?
Yes
Influence of external disturbance?
Yes
Exec 1- Vehicle controllability
No
Exec 2- Guidance defect
Overall failure?
Yes
Over 1- Loss of psycho-physiological ability
Over 2- Impairment of sensorimotor and cognitive abilities
Over 3- Exceeding cognitive abilities

HFF Method

Sheet N° 4: TYPICAL HFF GENERATING SCENARIOS (Top 30)

Category of HFF	Type of Human Functional Failure (HFF)	Typical Human Failure Generating Scenario
Detection	Detect 1 failure - Non-detection in visibility constraints conditions	'Detect 1C': Road user surprised by a pedestrian or a two-wheeler non-visible when approaching
		'Detect 1D': Driver surprised by the manoeuvre of a non-visible approaching vehicle
	Detect 2 failure - Information acquisition focused on a partial component of the situation	'Detect 2A': Focalisation on a directional problem
		'Detect 2B': Focalisation towards a source of information as a function of driver's layout representation
		'Detect 2C': Focalisation towards a source of information regarding the importance of the traffic flow
		'Detect 2D': Focalisation towards an identified source of danger
	Detect 3 failure - Cursory or hurried information acquisition	'Detect 3A': Cursory search for information while turning on the left (on the right for left driving countries)
		'Detect 3B': Cursory search for information while crossing intersection
	Detect 4 failure - Momentary interruption in information acquisition activity	'Detect 4A': Non-detection of the rapprochement from the vehicle ahead
	Detect 5 failure - Neglecting the need to search for information	'Detect 5A': Late detection of the slowing down of the vehicle ahead
		'Detect 5B': Late detection of a non-priority road user starting manoeuvre in intersection
Diagnosis	Diag 1 failure - Erroneous evaluation of a passing road difficulty	'Diag 1B': Under evaluation of the difficulty of an although known bend
		'Diag 1C': Erroneous evaluation of a bend difficulty in a context of playful-driving
	Diag 2 failure - Erroneous evaluation of the size of a gap	'Diag 2B': Erroneous evaluation of a merging gap connected to the low attention paid to the manoeuvre
	Diag 3 failure - Mistaken understanding of how a site functions	' Diag 3A': Mistaken understanding leading to a stopping failure in intersection

	Diag 4 failure - Mistaken understanding of another user's manoeuvre	' Diag 4B': Mistaken understanding of the other's manoeuvre related to the polysemy of their signals
		'Diag 4C': Mistaken understanding of other's manoeuvre related to cursory processing of the interaction
Prognosis	Prog 1 failure - Expecting another user not to perform a manoeuvre	'Prog 1A': Expecting a non-priority vehicle not to undertake a manoeuvre in intersection
	Prog 2 failure - Actively expecting another user to take regulating action	'Prog 2B': Erroneous expectation of the stopping of a non-priority vehicle approaching intersection
		'Prog 2C': Erroneous expectation of the stopping of a non-priority vehicle coming on the trajectory
	Prog 3 failure - Expecting no perturbation ahead	'Prog 3A': Expecting no vehicle ahead in a bend with no visibility
Decision	Dec 1 failure - Violation directed by the characteristics of the situation	'Dec 1A': Road user directed to go ahead in order to take the information
	Dec 2 failure - Deliberate violation of a safety rule	'Dec 2B': Overtaking on a zone with limited axial-visibility
	Dec 3 failure - Violation-error	'Dec 3B': Going ahead at intersection being drawn into manoeuvre
Execution	Exec 1 failure - Poor control of an external disruption	'Exec 1A': Sudden encounter of an external disruption
		'Exec 1B': Sudden encounter of an external disruption, more or less expectable
	Exec 2 failure - Guidance problem	'Exec 2A': Guidance interruption consequently to attention orientation towards a secondary task
		'Exec 2B': Guidance interruption consequently to attention impairment
Overall failure	Over 1 failure - Loss of psycho-physiological capacities	'Over 1A': Loss of psycho-physiological capacities consequently to a falling asleep or ill-health
	Over 2 failure - Alteration of sensorimotor and cognitive capacities	'Over 2A': Alteration of trajectory negotiation capacities
		'Over 2B': Alteration of guidance capacities
	Over 3 failure - Overstretching cognitive capacities	'Over 3A': Overstretching processing capacities in traffic interaction situation

HFF Method
Sheet N° 5: Involvement level of the driver

Coding Step: 6

CODE	Signification	Explanation
PC	Primary contributing	This modality designates the drivers who 'provoke the disturbance'.
SC	Secondary contributing	These drivers are not at the origin of the disturbance which precipitates the conflict, but they are however part of the genesis of the accident by not trying to resolve this conflict.
NC	No contributing	They are not considered as 'active' in the degradation because the information they had did not enable them to prevent the failure of others (contrary to the secondary contributing). They were not able to anticipate, due to this lack of information, the degradation of the situation, while the avoidance of the accident would have been possible in theory if this information had been supplied to them in time.
OP	Only present	These drivers are not involved in the destabilization of the situation even if they are nevertheless an integral part of the system. Their only role consists in being present and they cannot be considered as an engaging part in the disturbance.

HFF Method

Sheet N° 6: Emergency Failure

Coding Step: 7

Code	Description
Recovered	In the case when for the driver considered the avoidance manoeuver was adapted but the other one's neutralized this adaptation
ND	The AD of danger implies road users who did not detect the accident situation nor the emergency situation.
D1	The maneuver is the result of a decision forced by the situation constraints (offering no other choice).
D2	The choice of the maneuver that the road user decided to put forward is not suitable.
E1	The intention of maneuver is appropriate (adapted option) but the execution carried out is incorrectly because of several strong situational constraints.
E2	The intention of the performance is appropriate (adapted option) but not successful because of poor execution control issues.
Unavoidable	Distance / time conditions are too short to allow to achieve a successful avoidance.

HFF Method

Sheet N° 7: Crash Configuration

Coding Step: 9

For each vehicle and whatever the vehicle

	Code	Description
Primary crash configuration	**PCC.1**	Front
	PCC.2	Lateral
	PCC.3	Back
	PCC.4	roll-over
	PCC.5	Reversal
	PCC.6	Side swipe
	PCC.7	Unclassifiable
Primary crash Side	**PCS.F**	Front
	PCS.B	Back
	PCS.L	Left
	PCS.R	Right
	PCS.Ro	Roof
	PCS.U	Unclassifiable
Secondary crash configuration	**SCC.0**	No Secondary Choc
	SCC.1	Front
	SCC.2	Lateral
	SCC.3	Back
	SCC.4	roll-over
	SCC.5	Reversal
	SCC.6	Side swipe
	SCC.7	Unclassifiable
Secondary crash Side	**SCS.F**	Front
	SCS.B	Back
	SCS.L	Left
	SCS.R	Right
	SCS.Ro	Roof
	SCS.U	Unclassifiable

HFF Method

Sheet N° 8: Crash Aggravating Factors

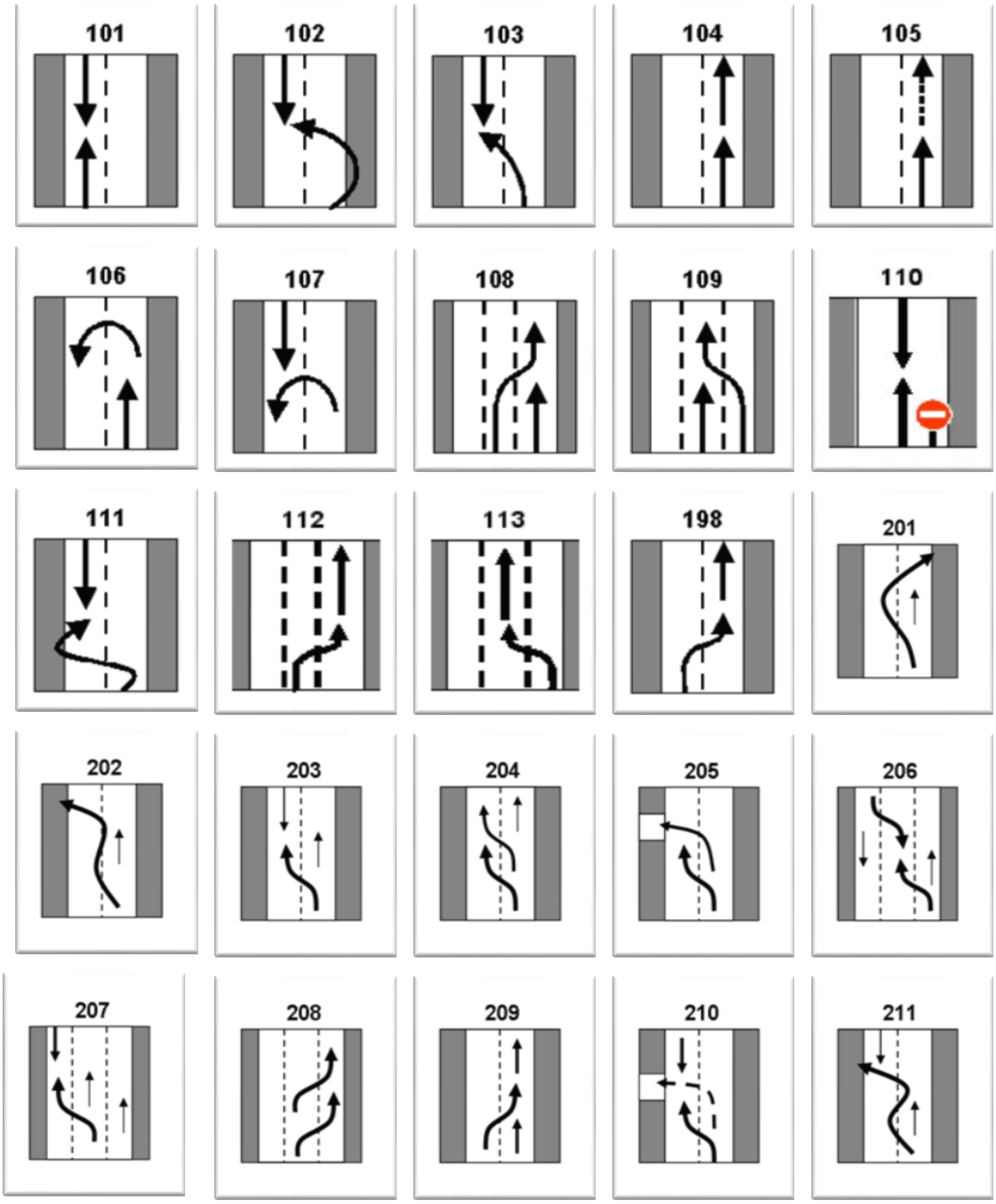
Coding Step: 10

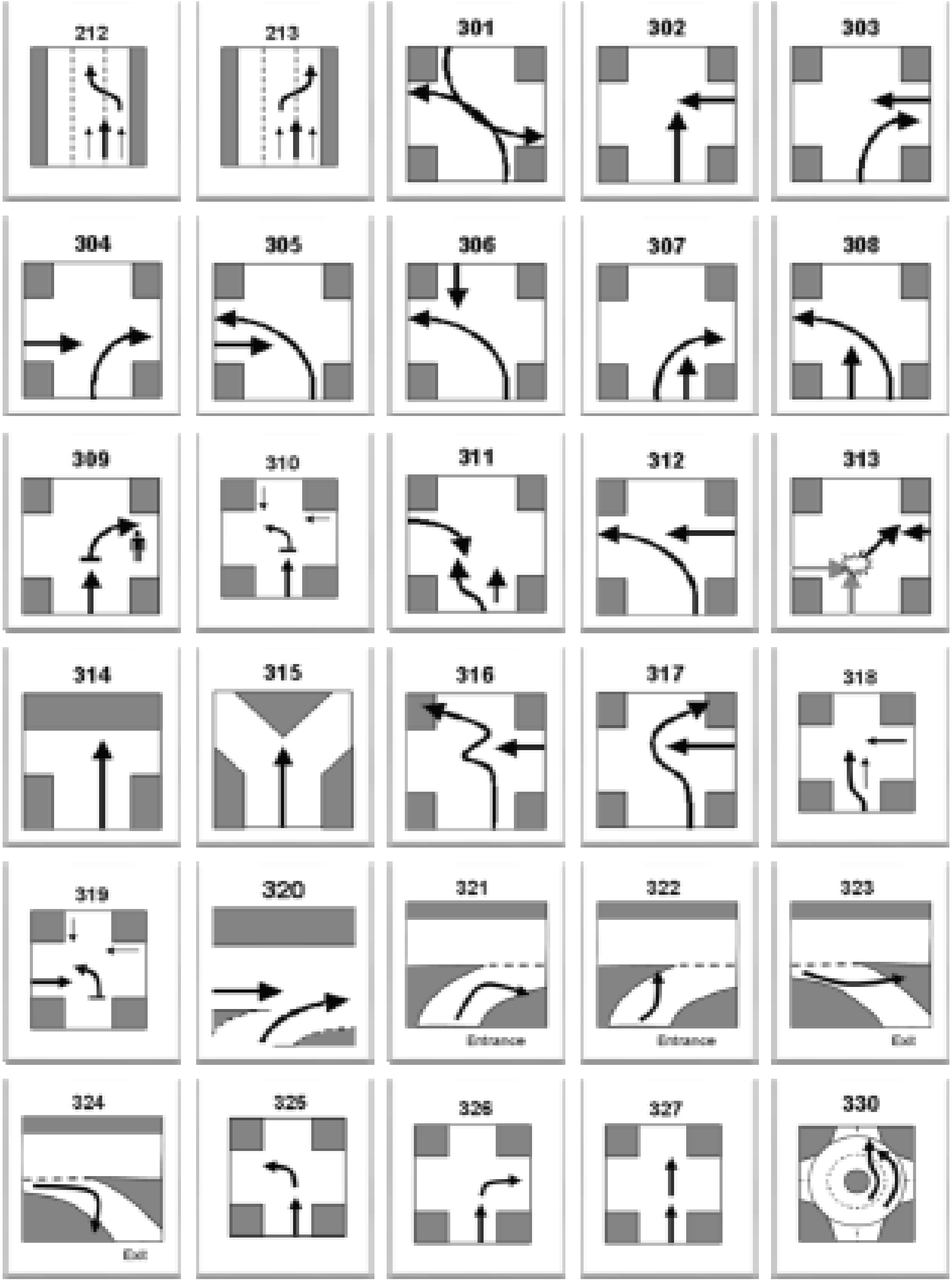
	Code	Crash Factors
Users related factors	**U.1**	Size of driver / passenger
	U.2	Weight of driver / passenger
	U.3	Old of driver/ passenger
	U.4	Gender of driver / passenger
	U.5	Medical condition
	U.6	Substances taken - alcohol
	U.7	Substances taken - Drugs
	U.8	Substances taken - medication
	U.9	Fatigue
	U.10	speed
	U.11	No braking
Main Impact type of main obstacle (having absorbed the greatest amount of impact energy)	**PC.1**	Nothing
	PC.2	small utility vehicle (minivan, derivative of a four-door sedan: cat 4)
	PC.3	small utility vehicle (van < 3,5 T : category 5)
	PC.4	off-road vehicle (category 7)
	PC.5	4X4
	PC.6	heavy truck (>3,5 T)
	PC.7	public transportation
	PC.8	Train
	PC.9	farm tractor
	PC.10	camper or small trailer
	PC.11	heavy construction vehicle
	PC.12	non-motorized two-wheel vehicle
	PC.13	motorized two-wheel vehicle
	PC.14	ground (only in the case of roll-over)
	PC.15	pole / lamp post
	PC.16	tree
	PC.17	guide rail
	PC.18	sign post
	PC.19	ditch - gutter
	PC.20	embankment
	PC.21	fence
	PC.22	wall / bridge pilon / building
	PC.23	lane divider wall

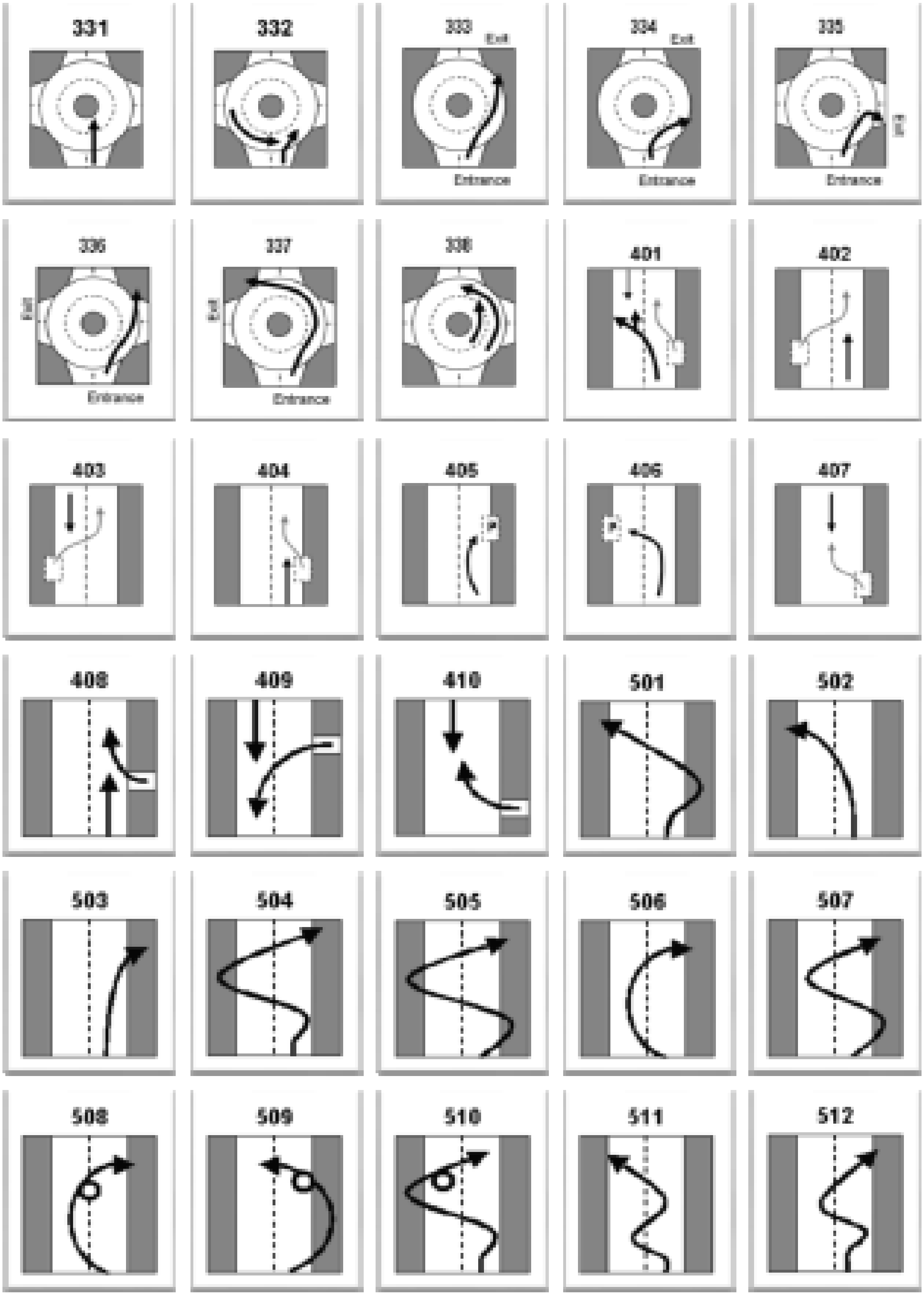
	PC.24	pedestrian
	PC.25	large animal
	PC.26	other
Secondary Impact (impact having absorbed less energy than main impact)	**SC.1**	Nothing
	SC.2	small utility vehicle (minivan, derivative of a four-door sedan: cat 4)
	SC.3	small utility vehicle (van < 3,5 T : category 5)
	SC.4	off-road vehicle (category 7)
	SC.5	4X4
	SC.6	heavy truck (>3,5 T)
	SC.7	public transportation
	SC.8	Train
	SC.9	farm tractor
	SC.10	camper or small trailer
	SC.11	heavy construction vehicle
	SC.12	non-motorized two-wheel vehicle
	SC.13	motorized two-wheel vehicle
	SC.14	ground (only in the case of roll-over)
	SC.15	pole / lamp post
	SC.16	tree
	SC.17	guide rail
	SC.18	sign post
	SC.19	ditch - gutter
	SC.20	embankment
	SC.21	fence
	SC.22	wall / bridge pilon / building
	SC.23	lane divider wall
	SC.24	pedestrian
	SC.25	large animal
	SC.26	other
Safety system	**SS.1**	Occupants completely ejected
	SS.2	Occupants partially ejected
	SS.3	Occupant pushed forward by thrust of rear occupant or load
	SS.4	Seat belt not available
	SS.5	Seat belt not fasted
	SS.6	Seat belt rear passenger unbelted
	SS.7	Seat belt rear passenger unavailable
	SS.8	child restraint system unbelted
	SS.9	Child restraint system properly fastened / defect
	SS.10	frontal Air-bag absent or defect
	SS.11	Lateral Air-bag absent or defect

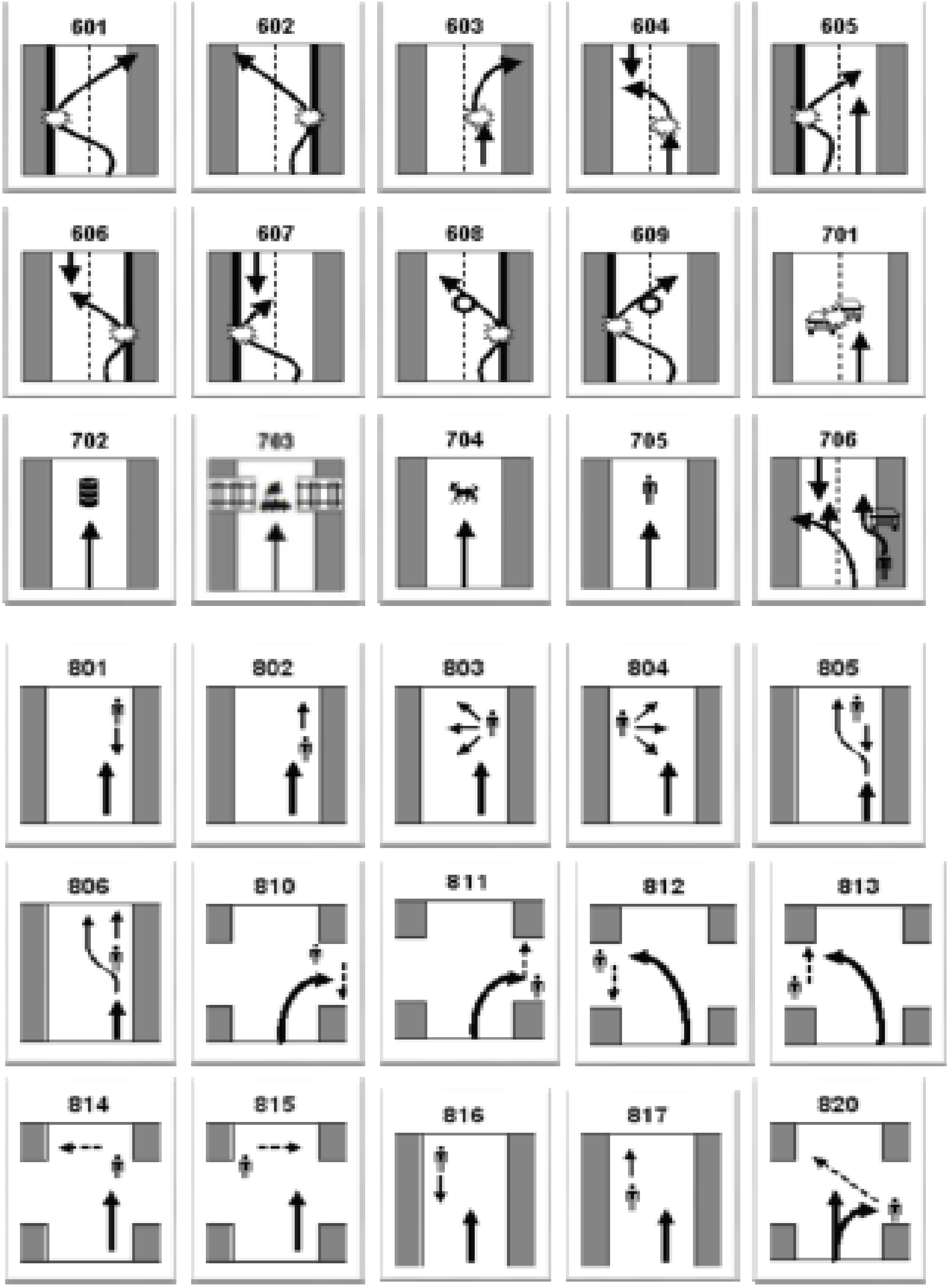
Sheet N° 9: Pictogram

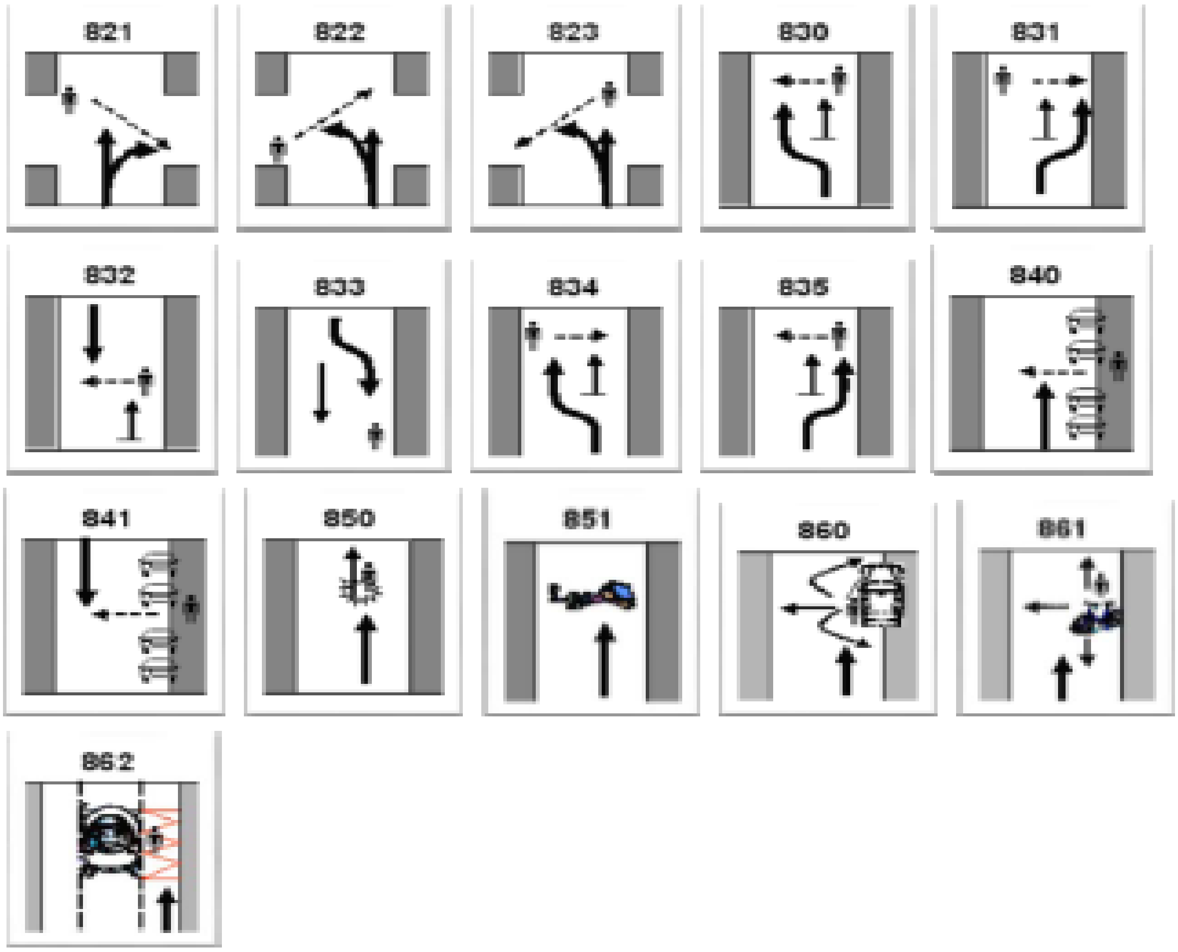
Coding Step: 11




212
213
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
Entrance
322
Entrance
323
Exit
324
Exit
325
326
327
330


331
332
333
Exit
Entrance
334
Exit
Entrance
335
Entrance
336
Entrance
337
Entrance
338
401
402
403
404
405
406
407
408
409
410
501
502
503
504
505
506
507
508
509
510
511
512


601 602 603 604 605
606 607 608 609 701
702 703 704 705 706
801 802 803 804 805
806 810 811 812 813
814 815 816 817 820


821
822
823
830
831
832
833
834
835
840
841
850
851
860
861
862