A User-Centered Approach to Road Design: Blending Distributed

Situation Awareness with Self-Explaining Roads

Miss Ipshita Chowdhury

Thesis submitted for fulfilment for the degree of Doctor of Philosophy

Institute of Infrastructure and Environment

School of the Built Environment

Heriot-Watt University

2014

ABSTRACT

Driving is a complex dynamic task. As the car driver drives along a route they have to adjust their driving technique in accordance with the traffic level, infrastructure and environment around them. The amount of information in the environment would be overwhelming were it not for the presence of stored mental templates, accumulated through training and experience, which become active when certain features are encountered. Problems occur when the environment triggers the incorrect templates, or fails to trigger the correct templates. Problems like these can be overcome by adopting a "self-explaining" (SER) approach to road design. That is to say, purposefully designed roads which trigger correct behaviour. A concept which can help improve the theoretical robustness of the SER approach is Situation Awareness (SA). SA describes how the environment and mental templates work together to ensure drivers remain coupled to the dynamics of their situation. It is a widely researched concept in the field of Human Factors but not in the domain of Self-Explaining Roads (SER), despite the very obvious conceptual overlaps. This thesis, for the first time, blends the two approaches, SA and SER, together. From this the ability to extract cognitively salient features and ability to enhance driving behaviour and their effects on driving behaviour are sufficiently enhanced.

After establishing SA as critical to driving through literature review the experiment phase started with determining the source of driver SA. Road environment was found to be of utmost importance for feeding into driver SA. This was also confirmed with the results of the on-road exploratory study. The success of the exploratory study led to large scale naturalistic study. It provided data on driver mental workload, subjective situation awareness, speed profile and endemic feature. Endemic features are unique characteristics of a road which make a road what it is. It was found that not all endemic features contribute to SA of a road system. Therefore through social network analysis list of cognitive salient features were derived. It is these cognitive salient features which hold compatible SA and facilitate SA transaction in a road system. These features were found to reduce speed variance among drivers on a road. The thesis ends by proposing a 'road drivability tool' which can predict potentially dangerous zones. Overall, the findings contribute to new imaginative ways road design in order to maximize safety and efficiency.

DEDICATION

To Mom, Dad and Road Safety in India

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

Appendix A Situation Awareness Rating Technique (SART)

Appendix B Driving Style Questionnaire (DSQ)

Appendix C Driver Declaration Form

Appendix D Participant Informed Consent

Appendix E NASA Task Load Index (NASA TLX)

Appendix F List of Publications

LIST OF FIGURES

LIST OF TABLES

Table of Acronyms

AASHTO: American Association of State Highway and Transportation Officials

ATC: Ait Traffic Control

BU: Bottom up

CARS: Crew Awareness Rating Scale

CAST: Coordinated Assessment of Situation Awareness of Teams

C-SAS: Canfield Situation Awareness Scale

CSF: Cognitive Salient Features

DSA: Distributed Situation Awareness

DSQ: Driving Style Questionnaire

DVLA: Driver and Vehicle Licensing Agency

ERC: Essential Recognizable Characteristics

ESRI: Environmental System's Research Institute

ETSC: European Transport road Safety Council

GDTA: Goal Directed Task Analysis

GPR: Global Prototype Routine

HF: Human Factors

HTA: Hierarchical Task Analysis

IEA: International Ergonomics Association

IRR: Inter-rater Reliability

LBFTS: Looked but failed to see

LSSR: Local State Specific Routine

MARS: Mission Awareness Rating Scale

MART: Malleable Attention Resource Theory

MOT: Ministry of Transport Test

NASA TLX: NASA Task Load Index

NDS: Naturalistic Driving Study

NHTSA: National Highway Traffic Safety Administration

NS: Naturalistic Study

PCT: Picture Sorting Task

PRT: Picture Rating Task

RDT: Road Drivability Tool

RLX: Rail level crossing

RPD: Recognition Primed Decision Making

S.D: Standard Deviation

SA: Situation Awareness

SABARS: Situation Awareness Behavioral Scale

SACRI: Situation Awareness Control Room Inventory

SACRI: Situation Awareness Control Room Inventory

SAGAT: Situation Awareness Global Assessment Technique

SALSA: Measuring Situation Awareness of Area Controllers within the Context of Automation

SARS: Situation Awareness Rating Scale

SART: Situation Awareness Rating Technique

SASHA: Situation Awareness for Solutions for Human-Automation Partnerships in European Air Traffic Monitoring

SER: Self Explaining Roads

SME: Subject Matter Experts

SNA: Social Network Analysis

SNDS: Semi Naturalistic Driving Study

SPAM: Situation Awareness Present Assessment Method

SWOV: Institute for Road Safety Research, Netherlands

TD: Top down

UCD: User Centred Design

UDRIVE: European Naturalistic Driving Study

Chapter 1 : INTRODUCTION

1.1 Introduction

It is true to say that the design and construction of road infrastructure is foremost the concern of civil engineers, but it is equally true to say that certain psychological parameters have been embedded in design guidance for a considerable number of years (AASHO, 2004; Highways Agency, 2011). As a result, the most modern form of road, the high speed, multi-lane, limited access motorway (the UK term for freeway, autostrada, autobahn etc.), provides drivers with what Vanderbilt (2008) describes rather fancifully as a 'toddlers view of the world':

"We make the driving environment as simple as possible, with smooth, wide roads marked by enormous signs and white lines that are purposely placed far apart to trick us into thinking we are not moving as fast as we are. [..] a landscape of outsized, brightly coloured objects and flashing lights, with harnesses and safety barriers that protect us as we exceed our own underdeveloped capabilities" (p. 90).

Motorways represent an unusual form of 'total engineered environment' and the behavioural outcomes are significant. In most countries in the world motorways are the safest types of road to travel on despite carrying by far the largest volumes of traffic (e.g. dft, 2008). Motorways, therefore, are an excellent example of a Self-Explaining Road, one that does not "need any additional explanation or learning process to know what it means and what to expect" (Stelling-Konczak et al., 2011; Theeuwes, 1995). The behavioural effects of the built environment, and the concept of Self-Explaining Roads (SER), suggests that a powerful mapping exists between the objective state of the built environment and the perceived state of that environment on the part of the individual within it. This concept is a familiar one. Norman (2002) refers to affordances and 'gulfs of evaluation', both of which describe a person's attempts to make sense of their context and how it matches their expectations and intentions.

Of course, it is not possible to create similar motorway-like 'total environments' in all situations, thus significant gulfs of evaluation can begin to occur on other, non-motorway types of road. SER research is about reducing such gulfs (Theeuwes & Godthelp, 1995) and the results so far are encouraging. This PhD is about extending this important strand of research by mapping the concept of situation awareness (SA) to Self-Explaining Roads (SER) so they can be designed from a novel user-centred perspective.

1.2 Background to the Research

1.2.1   Situation Awareness

Situation awareness is defined as, "activated knowledge for specific tasks, at specific times within a system (Stanton et al., 2006, p. 1291). Situation Awareness (SA) explains how drivers know 'what is going on (Endlsey, 199) and become coupled to the dynamics of their environment (Moray, 2004), SA is important. Prior research has shown it to be a greater determinant of accidents than improper speed or vehicle control (Gugerty, 1997).

SA is a popular and well-used term, and this thesis will be seeking to resolve some of the ambiguities involved in its use, as well as choosing an SA theory which maps on to the key characteristics of driving. Principle among these is that driving is a multi-agent system that is constantly changing. For the system to work there have to be SA transactions between elements of the wider driving system. For instance, on approaching a roundabout the driver checks which lane to take to reach their destination and also responds to traffic lights and other information. This involves a transaction of information between the road environment and the driver. While negotiating the roundabout the driver is very aware of their nearby lanes and constantly checks the mirror for other traffic. This involves a transaction of information between other traffic and the driver. While all this is happening, the vehicle will be communicating its speed and other parameters, like the extent to which it is rolling in the turn, the note of the engine, and so on. Thus there is a further transaction between the vehicle and the driver. These transactions are continuous and distributed across a system of interacting agents and actors. To a large extent, SA emerges out of these component interactions to give rise to the safe and efficient driving we wish to design into our infrastructure. In

this thesis it is argued that any system which is not able to support vital transactions between human or non-human elements in the driving context will cause performance reducing incompatibilities, with drivers becoming unsure of what to do and what is expected. The design goal, therefore, is to try and facilitate smooth SA transactions. This can be achieved if the driver / vehicle / road system undergoes design modification from human factors perspective.

There have been a large number of studies applying SA in road transportation contexts. These include exploring the SA of motorcyclists, pedestrians, car drivers and cyclists at rail level crossings (Salmon et al., 2013), the effect of secondary task performance on SA (Baumann et al., 2007), the association between driver distraction and SA (Young et al., 2013), the impact of advanced driver training on driver SA (Walker et al., 2009), the compatibility of SA between motorcyclists, car drivers, cyclists and pedestrians at urban intersections (Salmon et al., 2013), to name just a few. Importantly, though, none of the prior studies on SA have explored the connection between SA and SER and this is the gap this thesis operates within.

In a similar way, SER research has to-date been studied in isolation from SA. Since its inception in 1995 in the Netherlands by Theeuwes and Godthelp, SER researchers have mainly focussed on behaviourally relevant road categorisation (Weller et al., 2008, Theeuwes, 1998), speed reduction measures (Charlton et al., 2010), changes in road user behaviour following SER interventions (Mackie et al., 2013) and its relationship with forgiving roads and how it can improve the safety of rural roads (Herrsedt, 2006). SER research has tended to rely on simulator studies and simplistic psychological experiments. It is possible to go further.

Although the two concepts are currently disconnected, SER and SA both aim for the same outcome: to make roads which can interact seamlessly with their users. These are roads that are in line with driver expectations, where drivers know what to do and what to expect, where the easiest and most natural behaviour is also the safest and most appropriate one. The central argument put forward in this thesis is that this goal can be met in a new and effective way by coupling SA to SER.

1.2.2   Cognitive Compatibility

Some big effects on driver behaviour can emerge from comparatively simple design changes. Elliot (2003) reported a reduction of 1-2 mph just by changing some road signs and markings. Coloured road surfacing, and marking which create the appearance of narrowing of the road, led to a speed reduction of 5-7 mph. Speed bumps and horizontal deflections reduced traffic speeds by 10 mph (Elliot, 2003). Why is this?

Underneath the ideas captured by SA and SER is a further notion, that of 'cognitive compatibility'. Drivers have to be continuously aware of ever changing stimuli such as the shape and geometry of the road, the possibility of hazards, traffic lights, the state of nearby traffic, signs and signals and many more besides (Gugerty, 2011). The sheer quantity of information would be overwhelming were it not for the cognitive processes which allow drivers to perform in this environment. The driver is not merely a passive receiver and processor of information; they are an active participant in a wider information environment. In other words, skilled driving relies on inputs from the environment (things that are 'outside of the driver's head') but they also rely on knowledge structures and processes which reside 'inside the driver's head'. The key to the success of SER interventions is activating knowledge structures which give rise to the kinds of driving behaviours we want. This is not as easy as it seems. Malaterre (1990) attributed 59% of car crashes due to incorrect expectations of road users, in other words, features of the road environment which activated incorrect mental templates and subsequent behaviours. As such, one way to support drivers in forming correct expectations is by making a road predictable and recognizable (Konczak et al., 2010). Again, this is not as simple as it may appear. This is because the features engineered into road environments do not always give rise to the behaviours we wish, because the physical and engineering properties of roads are not necessarily the same as the cognitive properties. If these cognitive properties can be extracted and understood, they can be used to better activate knowledge structures and behaviours which we want.

Both SER and SA aim for an environment which needs "no further explanation or learning process to know what it means and what to expect" (Stelling-Konczak, 2011, p.102). This aim is achieved through affordances (Walker et al., 2013). Affordances are a key concept. Affordances are defined by Gibson (1977) as relationships which exist between people and the world around them as "actionable properties between the world

and actors" (Cited in Norman, 1999, p. 39). In everyday terms a straight, smooth, flat road 'affords' driving fast in the same way that a twisty, bumpy, hilly road 'affords' driving slowly. Combining SA with SER is, fundamentally, about discovering exactly what features 'afford' what behaviours, and designing them in to roads so that they are cognitively compatible with drivers.

1.2.3   Self-Explaining Roads

The idea of SER was originally funded by the European Commission 6th framework. As the name suggests, SER is about evoking safe driving behaviour through design (Kaptein & Claessens, 1998). SER researchers such as Kaptein et al. (1996) put forward the proposition that no amount or level of driver training can drastically reduce traffic accidents. In other words, simply telling drivers what to do is not adequate. Elicitng safe driver behaviour through design has been very successful in the past for speed reduction (Weller et al., 2008; Kaptein & Claessens, 1998; Charlton et al., 2010; Mackie et al., 2013; Martens, 2007). These studies have revealed that current methods for classifying and instructing drivers on correct speeds are not always in line with driver expectations. This leads drivers to choose an inappropriate speed. By applying SER treatments the researchers aimed to introduce 'inherent safety'.

In SER literature, the goal to change the characteristics of roads in order to influence driver behaviour is met with the identification of so-called endemic road features (Charlton et al., 2010). Endemic features include a wide range of entities and artefacts, from road width and lane markings through to landscaping and roadside furniture. These features make a road what it is via Essential Recognizable Characteristics (ERCs; Stelling-Konczak et al., 2011). For example, the presence of a 70 mph limit, blue sign boards, wide roads, smooth surfaces and very few two-wheelers are ERCs for a motorway. The review of SA shows us is that care should be exercised in how the identification of these features is undertaken, because ERCs can crop up across different road types and have different meanings. In other words, out of all the possible features in the built environment which ones are most 'cognitively salient'? The identification of these features in SER literature has proceeded along various lines and a good summary is provided in Charlton et al. (2010). Examples include several recent papers which use a form of 'picture sorting' task to discover what road features distinguish

different road types (e.g. Weller et al., 2008; Stelling-Konczak et al., 2011). Other methods include questionnaires (e.g. Goldenbeld & Van Schagen, 2007) and driving simulator studies (e.g. Aarts & Davidse, 2007). Curiously, very few studies take place in real road environments, and no studies have been identified which make meaningful reference to, or use of, SA as a concept.

1.3 Contribution to Knowledge

What we have are two closely related concepts, SA and SER, and prominent gaps in knowledge as to how they can be blended together in the name of inherent road safety, cognitive compatibility, and roads which contain all the 'right' cognitively salient features. From this the thesis seeks to generate three original and significant contributions to knowledge. They are:

The definition and extraction of cognitive salient features is the missing link between SER and SA research. As Charlton (2010) has reported more endemic features doesn't necessarily mean better safer roads or a better SER but a delicate balance is required between the two to make a road rich in SA and essentially a SER. This can be done by exploiting a recent theory of SA called Distributed SA (DSA). The contribution to knowledge is a link between a state-of-the-art theory of SA and SER, and the use of DSA inspired methods for extracting endemic features from real-world road environments.

Another substantial contribution to knowledge is proving that a system rich in SA derived 'cognitively salient features' (or affordances) will be more self- explaining than a road which does not incorporate these features. It should be possible to demonstrate that the preferred and normative speed should match on a defined SER compared to a non-SER.

The final contribution to knowledge is a tool or method to measure the SA embedded in the road. Coupling SER to defined SA methodologies should enable a Road Drivability Tool (RDT) to be developed through which road transport engineers can determine the 'cognitive properties' of routes, and target behavioural interventions more effectively.

1.5 Research Problem

1.5.1 Accident Toll

What is the significance of this problem? Why are SER a non-trivial topic worthy of study? According to the World Health Organization (2013) road accidents are a leading cause of death in the developed and developing world, with over 90% of those crudely attributed to human error (Olarte,2011). This conclusion is a familiar one. Human error has often been regarded as the major source of road accidents, but even though there is consensus that driving is a result of interactions between vehicles, environments and individuals researchers are often quick to blame the driver (Page & Elslande, 2007). After all, what is left from the immense progress made in the design of vehicles and road safety research?

Consensus does not always equate with probity, and some researchers, such as Stanton et al., (2002) have stated otherwise. According to Stanton (2009) humans have a tendency to blame themselves for all violations but research shows that unintentional speeding is often due to poor roadway design and signage placement (Salmon et al., In Press). The research problem lies in a simple idea. If an operator is unable to use the 'system' at their disposal successfully then the first conclusion to draw is that the device is not built for the user (Page & Elslande, 2007). Subjecting the driver to more and more training and educational programs will only mean efforts are being made on the driver to change, rather than other parts of the system. Although training is a common intervention with appealing common-sense logic it should be remembered that training is expensive, subject to degradation, and that not all people can be trained. It is not the easy option it is sometimes assumed to be. If education is not the panacea then what about engineering interventions? Engineering interventions might include features such as road signs and other overt sources of driver information. Unfortunately, road signs will only be read if the driver has an immediate need for the information (Herrstedt, 2014). The fact is, engineering, education and enforcement have been a three pronged strategy since the birth of modern road safety, yet we are still left with a substantial road death toll. How else can progress be made? It is argued that a new behavioural frontier needs to be opened up and exploited more proactively.

It is possible to reflect on the well- established guides, standards and references available to civil engineers wishing to design safe and usable roads.   These include the design manual for roads and bridges (DMRB) and the American policy on geometric design of highways and streets (AASHTO, 2004). Common to all such standards is a presentation of human factors principles and concepts, but at a very basic level. With ever-increasing pressures on transport infrastructure, new technology, and a persistent level of road deaths and casualties, existing guidelines appear increasingly basic. From a human factors / behavioural science point of view they do not equip the civil engineer with the required information and insights to design high performance 'user-centred' roads. The research presented in this thesis represents and attempt to address this.

1.6 Research Aims and Objectives

SA is a key theory underlying the thesis. As it stands currently, the dominant theory foregrounds the role of the individual, in that SA is viewed as existing inside the minds of individuals. As presented above, it is widely acknowledged that we are dealing with a road user 'system' and a systems approach to SA is needed in order to deliver the required insights. A new theory of SA, Distributed SA (DSA) is put to use to answer this concern.   Under a DSA viewpoint the SA of a system is maintained through interactions of human and non-human agents across the system, thus it is important to understand these interactions or 'SA transactions'. The dominant individual SA theory will tend to attribute accidents to individual factors, such as experience, abilities and training. Education and stricter laws is not the route to ensuring compatibility between drivers and their road environment. Thus, a paradigm shift is required from training and enforcement to cognitive compatibility via the provision of an SER environment.

Roads should be designed to make it easy for the road user to go right, and difficult to go wrong. Coupling SER to DSA enables progress to be made on this front. In this respect the thesis attempts to lay the foundation for user centred roads through DSA.

SER lacks a clear methodology. The concept of endemic features put forward by Charlton et al., 2010 also lacks validation from the SER community. A survey of methods reveals there are no practical tools for assessing the behaviour change in

different sections of the road. Since SER is a new and very specific concept research has been limited compared to SA. Through years of research the DSA methodology is found to be scientifically robust and has been applied successfully in various domains (Walker et al., 2013, Salmon, 2008, Stanton et al., 2013). Coupling DSA to SER gives the latter access to the former's rigour in method and approach, enabling the extraction of cognitively salient features and validated by real drivers.

The underlying thesis is that if we consider the road transport system as a whole then it can be designed as a whole to help drivers generate optimum levels of SA and, in turn, support safer driver behaviour. This can be practically achieved if the existing design guidelines which, from a human factors perspective are comparatively basic undergo modifications from an SA perspective. The **objectives** defined for this research are to:

1. Explore different aspects of situation awareness and select a model most suited to of the systemic properties of road infrastructure.
2. Determine ideal source of driver SA and compare how driver's information needs match the results gained from real world data collection.
3. Understand the self-explaining characteristic of different British roads and reveal the link between SER and SA.
4. Extract cognitively salient 'endemic features' required by SER approaches using DSA methods and validate them.
5. Develop a tool to assess the cognitive compatibility of roads and validate it against real-world accident data.

1.7 Thesis Structure

To answer the research questions posed, and meet the aims and objectives, the thesis proceeds through a number of defined steps. These include reviews of the knowledge-base and literature, reviews of SA methods, pilot studies exploring the mapping between SA and SER, a full scale naturalistic driving study, a study validating the extracted endemic features to see the extent to which they are implicated in actual driver behaviours, and finally, the development of a route driveability tool. A brief description of the Chapter contents is provided below, along with a table showing how the Chapters map on to the main objectives:

Table 1.1 Objective-wise Structure of Research

Aim	Objectives	Methods
To lay the groundwork for developing user centred roads using the existing research on Situation Awareness and Self-Explaining Roads. roads using the existing research on situation awareness and self-explaining roads.	Objective 1: Explore different aspects of situation awareness and select a model most suited to of the systemic properties of road infrastructure.	Literature Review Review of Methods (Chapter 2 and Chapter 3)
	Objective 2: Determine ideal source of driver SA and compare how driver's information needs match the results gained from real world data collection.	Situation Awareness Requirement Analysis Exploratory naturalistic study (Chapter 4 and Chapter 5)
	Objective 3: Understand the self-explaining characteristic of different British roads and reveal the link between SER and SA	Pilot Study (Chapter 4)
	Objective 4: Extract cognitively salient 'endemic features' required by SER approaches using DSA methods and validate them	Naturalistic Driving study (Chapter 5,6,7)
	Objective 5: Develop a tool to assess the cognitive compatibility of roads and validate it against real-world accident data	Picture Rating Task Route Driveability Tool (Chapter 8 and 9)

Chapter 2: Situation Awareness: Literature Review

Chapter 2 provides a comprehensive picture of the current state-of-the-art in SA theories. It explores related concepts such as mental models, schemas, distributed cognition and the perception cycle. The chapter concludes with a series of definitions for SA and fulfils the first objective (refer to table 1.1) of this Thesis.

Chapter 3: Review of Measurement of Situation Awareness Techniques

The third Chapter explores how SA can be measured. A detailed review of over 20 methods for measuring SA is conducted, compared and contrasted to arrive at a suitable method to fulfil the aims and objectives of the thesis. Chapter 3 develops the methodological framework to be adopted for the rest of this thesis. Together with chapter 2 it also contributes to fulfilling Objective 1.

Chapter 4: Does Situation Awareness map onto Self Explaining Roads

This chapter marks the beginning of the experimental Chapters in the thesis. Chapter 3 sets out to determine what is expected by the driver from the environment they are driving within and what is actually provided by the road. Two sets of experimental data are analysed to achieve Objective 2 and a trade-off between driver expectation and present road design is spotted. This Chapter consolidates the missing link between SA and SER, which is the underlying aim throughout this thesis.

Chapter 5: Semi-Naturalistic Driving Study Method (SNDS)

This Chapter details the setup of conducting a large scale naturalistic driving study. It contains the process from recruiting participants to the materials used, route selected and actual conduct of the experiment. In essence, Chapter 5 uses the results of Chapter 4 in a broader scale thus contributing towards the method's validity.

Chapter 6: Direct and Indirect Measures of SA and Driver Behaviour

This Chapter presents driving performance, workload and self-assessed SA data stemming from the naturalistic driving study. It goes onto the study of speed profiles of the participating drivers in the two roads, and points out how they map with the already placed speed limits. Chapter 6, along with contributing to the wider aim of the thesis, also highlights some limitations with the most widely used methods for SA assessments.

Chapter 7: Extracting Endemic Road Features

Chapter 7 establishes the bridge between SER and SA through DSA network analysis approaches. These are used to extract endemic road features in a highly novel way. It meets Objective 4 and draws out design elements which are more important to drivers than others. This chapter also explains the SA transactions between different agents in the road and how this affects driver behaviour.

Chapter 8: Validating the Distributed Situation Awareness approach to Endemic Road Features

The features extracted Chapter 7 are formally tested for their reliability and validity. The behaviour change of drivers with and without endemic features is revealed in Chapter 8. It also explores how these features can contribute to road safety through previous accident data literature.

Chapter 9: Road Drivability Tool

Chapter 9 brings together the findings into a practical tool. This chapter describes how the drivability of a road can be determined using a robust, theoretically derived step-by-step method. .

Chapter 10: Conclusions and Future Research

Chapter 10 draws out the conclusions on designing a user centred road and remarks on its feasibility, wider applications and contribution to the existing body of knowledge. The chapter later analyses the limitations of the Thesis and some possible areas to be looked into in the near future.

Chapter 2 : Situation Awareness: Literature Review

2.1   Introduction

Self-explaining roads (SER) are roads environments designed to elicit a correct assessment from road users as to what constitutes appropriate driver behaviour (Macharis et al., 2005). In essence SER are developed to increase the inherent safety of a road by taking into consideration human perception and information processing abilities (Weller et al., 2011). SA is very relevant to a SER approach. SA explains how drivers use information from the world (road environment) to combine long-term goals with short-term goals in real time (Sukthankar, 1997). SA and SER rely on affordances (e.g. Weller et al., 2010), schemas (e.g. Charlton et al., 2010) and expectation and prediction (Stelling-Konczak et al., 2011). Furthermore, SA is inherently a safety-related concept (Salmon & Stanton, 2013) which makes it even more fitting to be studied in the context of SER.

SA is one of the most popular Human Factors and Ergonomics concepts studied (Patrick & Morgan, 2010). Consequently, SA is backed with much research making it theoretically rich. The concept of SA remains highly contentious until this day (Stanton, 2010) and much debate exists over its theoretical understanding. Therefore to link the SA with SER the first step is to determine which SA model is appropriate for this thesis and what school of thought is most suitable to explain it in terms of SER. Hence, Chapter 2 is written to accomplish the first objective: "explore different aspects of situation awareness and select a model most suited to the systemic properties of road infrastructure" The chapter begins with exploring SA as a construct and following that explains three different schools of thought of SA. After selecting one SA school the theoretical basis of each model is described, detailing its positives and negatives. Finally, this chapter discusses the plethora of definitions of SA and ends with deriving conclusions from the SA literature reviewed.

2.2 What is Situation Awareness (SA)?

Usually, SA has been considered to be intuitive (Sandom, 2000). This means situation awareness has been equated to intuition, which is defined as the ability to acquire knowledge without inference or the use of scientific evidence (Oxford English Dictionary). Intuition, the popular Jungian psychology term, is based on merely beliefs and cannot be justified logically in a world based on sound reasoning and science. Though considered intuitive in the beginning, it has now been fully proven that SA is a scientific construct developed after decades of research on mental models, schemas and advanced cognition. If not intuitive, SA was considered as common sense (Stanton & Piggot, 2001). Common sense provides defaults and general plausible inferences that aid humans in understanding the world and their actions (Freedman & Adams, 2009). Surprisingly, Freedman & Adams (2009) proposed in a technical report that SA can be maintained through common-sense reasoning. Like intuition, the notion of SA to be common sense was totally negated with advancement in research on system ergonomics. Systems ergonomics considered SA to be resident in the system, of which humans are a part. System and sub-system are not expected nor proven to have any common sense.

SA as a construct remains a highly popular area of research within the ergonomics community. It is a concept that describes how operators such as drivers in complex systems like vehicles, roads, etc. develop and maintain awareness of 'what is going on' (Endsley, 1995). Particular interest has been taken by the aviation industry, however; similar complex systems have started to realize the potential of SA (Harris, 1997; Garland & Endsley, 1995). Sarter and Woods (1991) identify SA to be a critical but ill-defined phenomenon in complex dynamic systems. SA is also identified as a prerequisite for operating safely in any complex dynamic system (Sarter & Woods, 1991).

Research in the last decade has showed that incorrect or faulty SA is the causal factor in the leading of aviation mishaps (Hart et al., 1991). Endsley (1999) also argued that 88% of human error can be attributed to faulty SA. Treat et al. (1979) reported that improper lookout and inattention were two major causes of road accidents in the state of Indiana, USA. Gugerty (1997) suggested inattention and improper lookout to be aspects of SA.

This means SA is the major cause of road accidents. Additionally, Gugerty (1997) reported that poor SA contributes more to accidents than incorrect decision making (improper speed) or psychomotor ability (improper driving technique). Thus, maintaining a high level of SA has become more of a compulsion than a choice (Endsley, 1997). Certainly, based on the challenges SA poses and the significance it has established, maintaining high levels of SA is very demanding (Endsley, 1997) yet critical for safety. Definitely, SA has come a long way since its inception in the 1990s. From being restricted to fighter plane crash analysis, SA principles are now applied to system design. The existence of SA as a stand-alone concept (Wickens, 1992; Dekker & Hollangel, 2004) was questioned but it has now branched out into three different perspectives (Sorensen et al., 2011). Surprisingly, Wickens (2008) has acknowledged the contribution of SA in safety critical domains, after initially questioning the very existence of SA in 1992.

Situation awareness has stemmed out of aviation but now it has spread its roots to various domains. Domains which are commonly the focus of SA-related applications include aviation (Endsley, 1999), military (Stanton et al., 2010), air traffic control (Kaber & Endsley, 2004), rail (Golightly et al., 2010), process control (Patrick & Morgan, 2010) and healthcare (Hazlehurst et al., 2007). Furthermore, Salmon (2008) in his review of 20 peer-reviewed academic journals identified diverse areas of SA application, ranging from human factors and ergonomics to sports science to computer graphics, artificial intelligence and disaster management.

Another domain where SA has played an important role is road transport, which is of particular relevance to this thesis. Some researchers such as Young et al., 2011 reported that SA has received less attention in a road transportation context despite being highly relevant to it. The application of SA in surface transportation is triggered by pioneering results from the studies of Gugerty (1997) and Treat et al. (1979) which put forward that failure related to poor SA is a major cause of road accidents, rather than inattention. Poor SA of the driver can result in inattention (Salmon et al., 2011; Dingus et al., 2006) and misjudgement – that is, incorrect comprehension of the present situation leading to erroneous projection of their vehicle in a real-time environment (Endsley,1995). Brumec et al. (2010) who proposed road safety as a factor in the road environment, cautioned not to ignore the contemporary findings of SA while discussing road safety.

Allen and Hill (1982) found supplemental traffic signalling results in the greatest degradation of driver performance, as well as the highest driver response time. Thus, this directly increases the complexity of the system which is a contributor to poor SA of the system. Furthermore, this study of traffic signal display complexity also indicated that the traffic signal itself has poor SA, which transfers incorrect or poor SA across the system, and thus reduces driver performance. These findings make SA imperative to study in the road transport domain.

Based on the notion that appropriate conceptualisation and assessment of SA has much to offer to future road safety efforts, this research focuses on the selection of the most suitable model of SA and SA methods' to understand the existing road scenario, and additionally, to explore the possibility for constructing "SA rich roads."

2.2   Origin of Situation Awareness

SA is an interestingly ubiquitous concept which makes it difficult to pinpoint its exact source of origin. A group of ergonomists date the origin of SA to World War II (Endsley, 1995). The term was used by members of United States Air Force in the context of air-to-air combat during the Vietnam War (Watt, 2004). Another group (Stanton et al., 2001) argues SA to have its roots in World War I. In this context, SA was described as a problem in perception and Stanton et al. (2001) later confirmed perception to be the main essence of the definition of SA. This construct of SA was simply lost for decades and often referred to as "common sense" (Stanton et al., 2001).

The influential term for the level of individual awareness of the situation officially called SA, didn't come into the limelight until 37th volume of the journal *Human Factors* dedicated to situation awareness itself. Mica.R.Endsley provided a pioneering explanation for SA in that volume, strengthening this definition by providing a linear 3-level model as its theoretical underpinning. Undoubtedly the pioneers of this field are Endsley and colleagues, as they were the first to develop a scientific explanation of the construct. Since then, the concept of SA has grown enormously in the last 50 years and has had a major influence in practice and research in human factors (Patrick, 2009). Road safety however benefited from this growth fairly recently.

The growth of SA as a vital concept occurred very rapidly and has been instantiated by the results of the study done by Patrick & Morgan (2009) at Cardiff University. After putting SA or situation awareness in Google and Google scholar on 27/11/2008, 203,000 hits were obtained from the former and 17,500 were retrieved from the latter. A more specific search in PsycINFO where SA was used as a keyword or a title of journal found an increasing number of articles with a sharp increase from 1995, shown in Figure 2.1 (Patrick & Morgan, 2009). The results not only reiterate the importance of SA as a construct but also indicate a rise in the research of the construct after 1995.
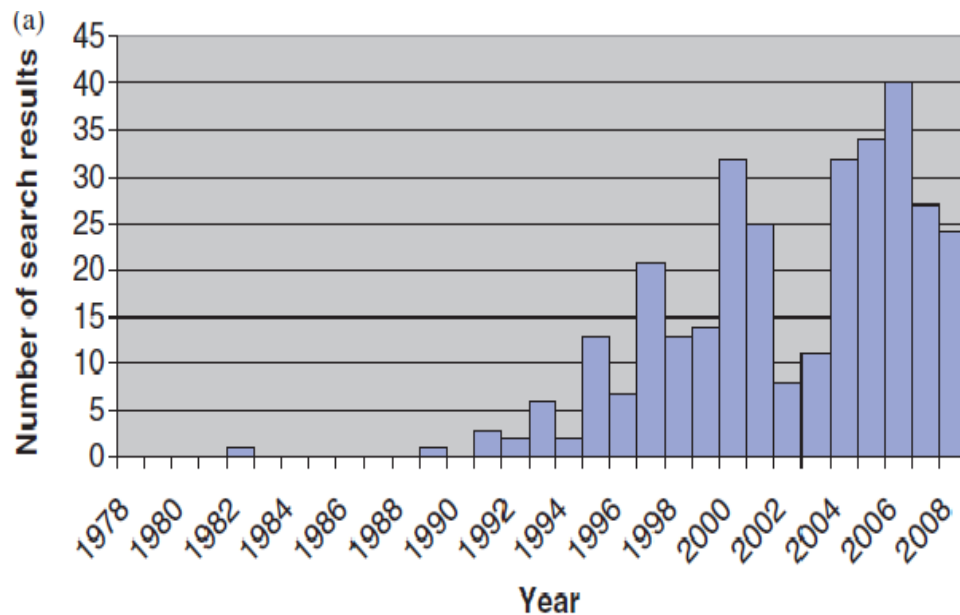
Figure 2.1. Research Trend in Situation Awareness (Patrick & Morgan, 2009)

In summary, SA has developed as a widely-applicable concept such that it is difficult to contain it within one single discipline of study. It has permeated research in engineering, computing, psychology, and many other related disciplines. Apart from the 37th volume of *Human Factors* there have been numerous issues in journals such as *Ergonomics* and *Theoretical Issues in Ergonomics Science* where SA has been dissected and explained. This further affirms the importance SA has gained as a construct.

2.3 Three Schools of Thought on Situation Awareness

Table 2.1. Perspectives on Situation Awareness

Perspective on Situation Awareness	Example Models	Unit of Analysis in Road Transport Context
Individual or Psychological	Three-level model (Endsley, 1995)	Drivers
Computing or Engineering	Data flow model (Diaper, 2004)	Information sources for road users, eg, displays, road signage, in-vehicle systems
Systems and Human Factors	Distributed situation awareness (Salmon et al., 2009)	Overall road transport system and interactions between human (e.g. drivers, pedestrians, bicyclists) and non-human agents (infrastructure, vehicles, signage, etc)

The three schools of thought regarding SA are very contrasting in nature and often contradict each other (Sorensen & Stanton, 2011). The vast difference amongst these schools has hindered the development of a universal definition for SA. Various models of SA have been presented in the academic literature, relating to the level of SA held by individuals (Endsley, 1995), teams (Salas et al., 1995) and socio-technical systems (Salmon et al., 2009). Stanton et al. (2010) referred to these aforementioned perspectives as individual/psychological (for Endsley), computing/engineering (for Salas) and systems and human factors perspective (Salmon et al., 2009).

There is considerable debate and competition amongst these schools of thought and approaches. The need to select an appropriate SA perspective has become more important now than it was before. The advent of the "designing for SA" concept (Endsley & Jones, 2001) has made it inevitable to follow a school of thought of SA. Notable contention is found whilst explaining the working or development of situation

amongst different agents in teams (Salmon et al., 2009; Salmon et al., 2011). Sandom (2002) classified SA under two perspectives, namely cognitive and interactionist. The cognitive perspective is the most prevalent view of SA and puts forward that SA is 'in the head of the operator' (Sandom, 2000). Cognitive perspective theorists often confusingly refer to SA as a cognitive process, a state of knowledge or both (Sandom, 2000). With this distinction, *product* refers to the state of awareness, whereas *process* refers to the various cognitive activities involved in acquiring and maintaining SA (Sandom, 2000). Contrary to the cognitive perspective, the interactionist perspective is of the view that SA is an abstract concept which is resident within the interaction between the individual and the environment (Sandom, 2000). Despite this fundamental divergence between the two perspectives there is some conceptual similarity. Both of them acknowledge the importance of SA in safety and this can be manifested through designing for SA. This conformity forms one of the thrusts for this thesis and can be seen in Figure 2.2.
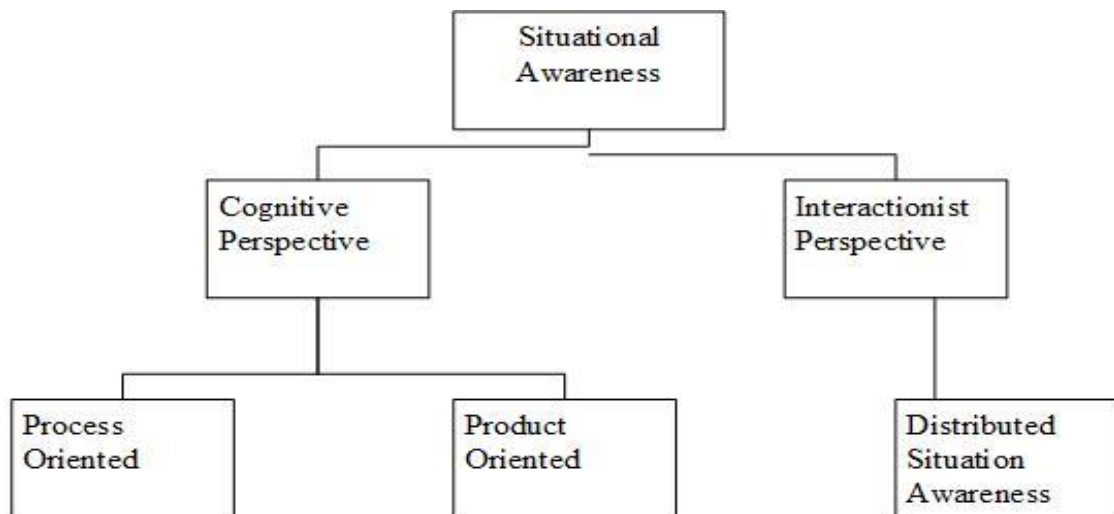
Figure 2.2. Dominant Perspectives of SA Adapted from Sandom (2000)

From this review of extensive literature from reliable sources, a number of conceptual issues will be identified which form the basis of such debates between various approaches of explaining situation awareness. Innate to this debate of approaches is another age-long debate of SA being a product or process. All of these will be explored in the subsequent sections.

2.3.1. Psychological/ individual perspective

The individual/psychological perspective argues that SA is purely an internal cognitive construct couched inside the heads of individuals (Salmon, 2008). According to this, SA exists only in the working memory of human operators. The models under this perspective are often referred to as foundational (Salmon, 2008) in nature as they led to development of the many other models of SA which are now used. The most profound definition of SA (2395 Google Scholar citations) comes into this category and defines SA as a product:

"Situation awareness is the perception of environmental elements with respect to time and space and comprehension of their meaning and projection of status after some variable has changed, such as time" (Endsley, 1995, p.5).

All the definitions from this perspective point to SA as being aware of what is happening in the vicinity to understand how information, events, and one's own actions will impact goals and objectives both immediately and in the near future (Adams et al., 1995; Smith & Hancock, 1995; Endsley et al., 1995; Billings, 1995; Carol, 1992; Haines & Flateau, 1992). The pioneering model under this perspective is Endsley's three-level model. The three-level model described by Endsley (1995) explains SA as the end product of skimming information through the three levels. Please refer to figure 2.3 in the next page. SA consists of three separate levels: perception, comprehension and projection. Level 1 is the perception of elements in environment. Level 2 is the comprehension of the meaning of those elements. Level 3 is using the knowledge from Level 1 and Level 2 to project the future states of a system (Endsley, 1995).
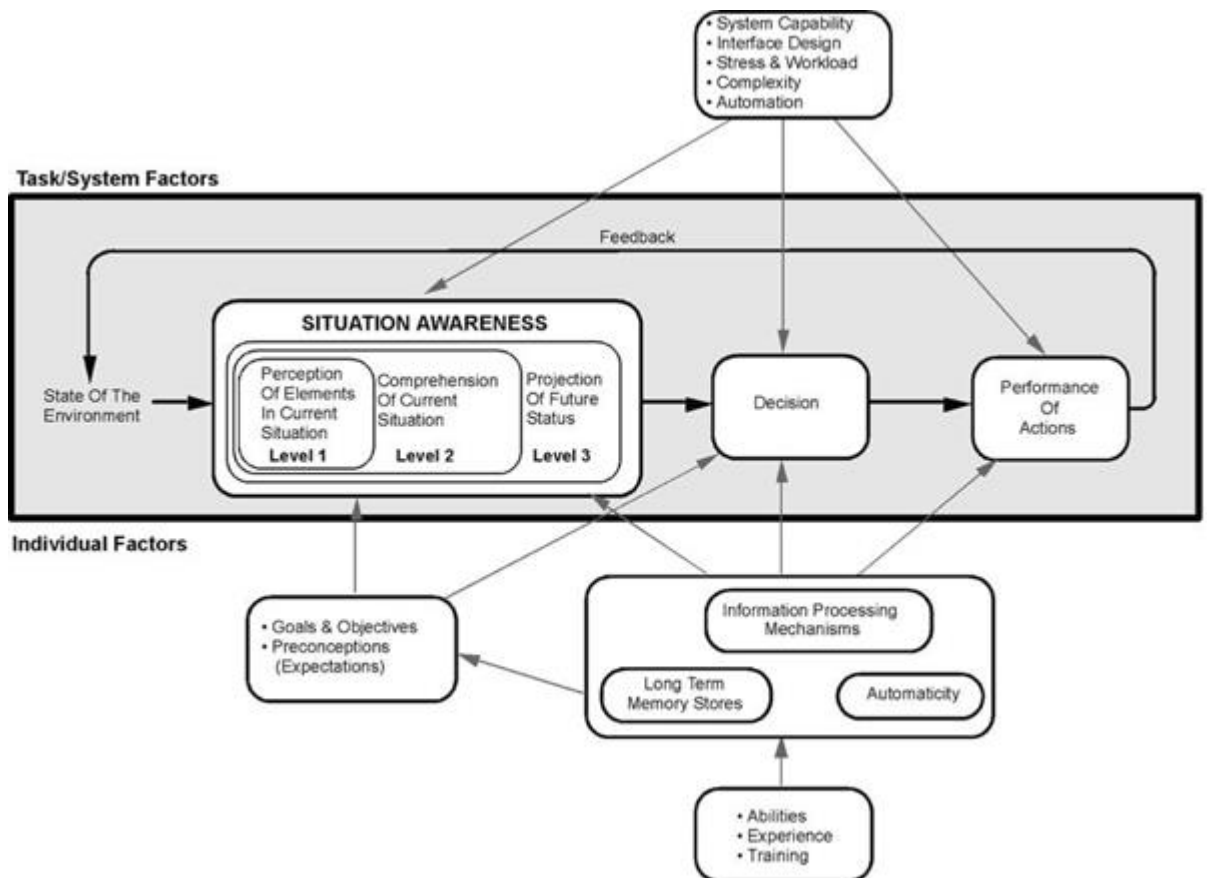
Figure 2.3. Endsley's 3-Level Model (Endsley, 1995)

It is evident from Figure 2.3 that SA is formed after a sequential flow of information from one level to another. Without meaningful and adequate comprehension (Level 2) of Level 1 (perception), it is not possible to make assumptions or project (Level 3) the future of the state of the environment (Banks et al., 2011; Sorensen & Stanton, 2011). In other words, it is essential to form a representation of a situation before a decision is made. In a road transport context, Endsley's three-level models argues that driving efficiency can be improved by focussing on individual road user interventions such as advanced driver training (Salmon, Stanton, & Young, 2011) and any form of feedback from road infrastructure or in-vehicle assistance system should strive to enhance the three levels of SA.

Isaac (1997) used Endsley's model to explain SA formation in air traffic control, maintaining that the ability to produce a mental representation of a situation enables an air traffic controller (ATC) to regain and preserve SA. Isaac (1997) proposed a model (shown in Figure 2.4) that depicts an explicit link between SA and the production and use of mental models using Endsley's three-level model.
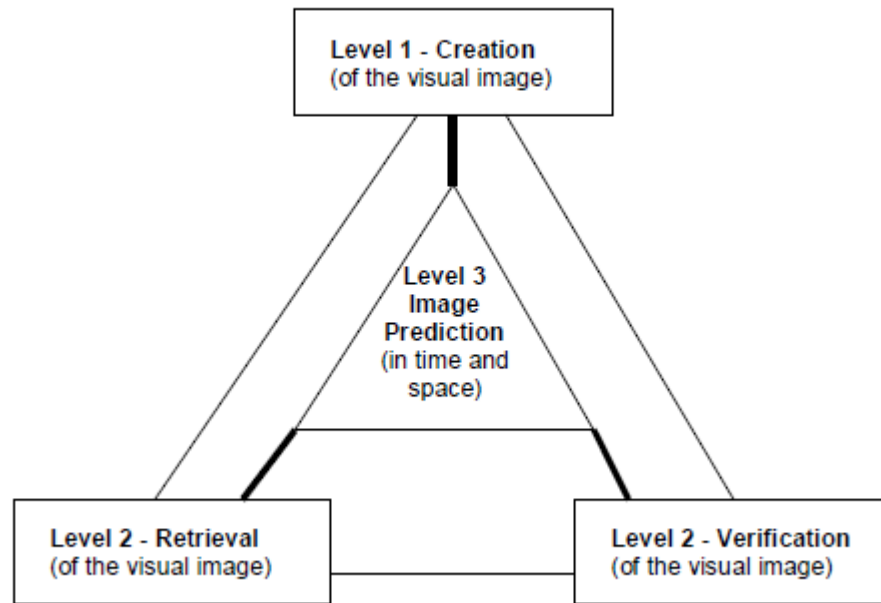
Figure 2.4. Isaac's Adaptation of the 3-Level Model (Isaac, 1997)

Central to the three-level model is the assumption of existence of mental models, and the role they play in the formation of SA as a product. A mental model is defined as a dynamic mental representation of the situation that allows people to predict future states and infer situations which have not been experienced by the operators before (Woods et al., 1994). The details of mental model formation will be discussed in the theoretical underpinning section.

This approach emphasises the importance of psychological properties of an individual to achieve SA. Sarter and Woods (1995) regard SA to be an array of cognitive processing activities that are critical to perform in a dynamic system (Sorensen & Stanton, 2011; Salmon, 2008). According to Banbury et al. (2004) the individual is able to perform this range of activities by developing a mental picture or theory of the world that aids in the process of how the elements are to be understood. Another group of researchers like Bedny and Meister (1999) and Artman (2000) argue that SA phenomena can be understood as a part of cognitive activity that is dynamic in nature. For example, Bedny & Meister (2000), describe the formation of SA as a phenomenon manifested though activity approach that outlines various cognitive processes that are associated with human behaviour (Stanton et al., 2010) rather than following a mental model. Refer to figure 2.5. The theory of activity (Bedny & Meister, 1999) asserts that

individuals have certain goals in the form of desired end of the activity, and a plan or methods to accomplish them. Certainly, this plan is motivated and inspired by the current state of the system. The difference in the actual state and desired end of the activity makes an individual take action towards the goal. Bedny and Meister (1999) suggest that each of the blocks (white, numbered) form a functional role in development and maintenance of SA, and if required orient themselves for the achievement of SA (Salmon et al., 2009). This again re-establishes Artman's (2000) definition of SA as 'active construction of situation model.'

Thus, it is very apparent that the sole responsibility for maintaining and developing SA lies with the individual only, which consequently, increases the mental workload of the driver. To summarize, the individual school of thought takes into account several cognitive processes such as memory, information processing, decision making, time, etc. that contribute to the development of SA inside the head of the operator. The most important processes as underlined by Stanton (2010) include attention, memory, experience, goal-driven experience, etc. Clearly, so many cognitive factors forming SA proves to be a limiting factor (Endsley, 1995; Endsley, 1999; Endsley, 2000; Banks & Millward, 2009). Some of these limitations will be dealt with in the section related to definition.
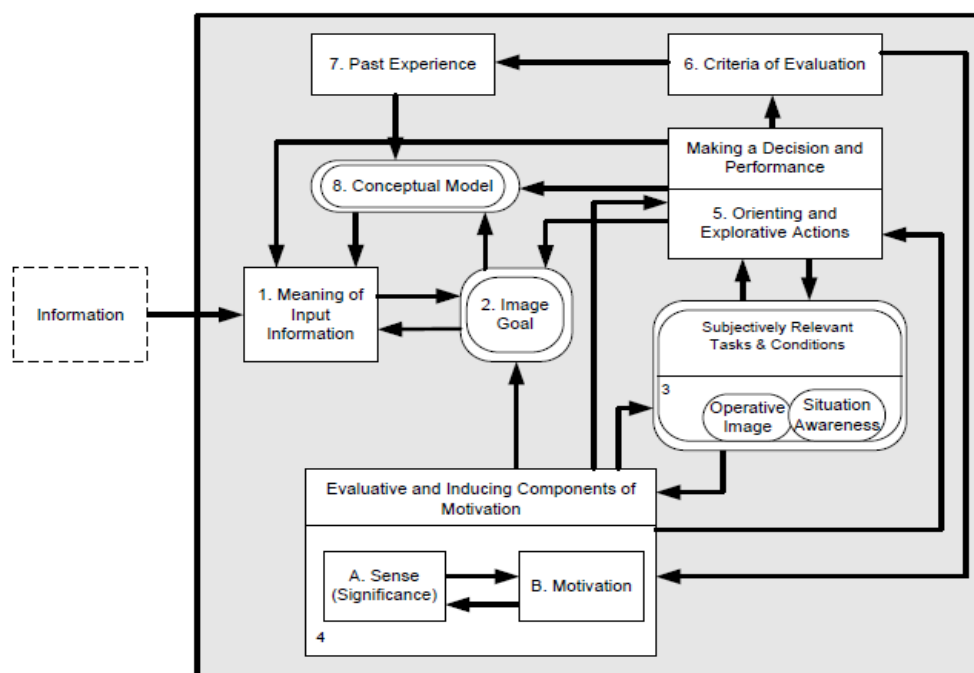
Figure 2.5. Activity Theory (Bedny & Mister 1999)

2.3.2. Engineering or computing perspective

The computing or engineering perspective is dominant in the military including all branches of the armed domain (Ackerman, 2005). The computing/engineering perspective presents a 'world view' towards SA (Stanton et al., 2010). SA is identified as a physical phenomenon (Stanton et al., 2010; Stanton et al., 2011; and Salmon et al., 2011) which is in the artefacts of the environment rather than the individual. Hence, the artefacts are the SA holders (Stanton et al., 2010). This adopts the view that the world excludes the individual, so the operator does not have many tasks to do regarding SA.

These views from the engineering or computing perspective are in stark contrast to the psychological or individual perspective of SA. The main difference is that this approach purports that the artefact itself is the holder of SA rather than the individual. Jenkins et al. (2008) reported that pilots refer to their displays as containing their SA. Pilots are required to mark their route of flying before commencing the flight; they do so by marking the route with rivers, power lines and crutches. Jenkins et al. (2008) confirmed that it is these symbol of rivers, etc. that have the pilot's SA. A similar phenomenon is also seen in road transportation. The car drivers are required to set the route on their GPS before actually starting to drive on the path defined. Thus, translating the results from the aviation study by Jenkins et al., 2008 it will be wise to assume that traffic lights, roundabouts, road width, etc. have the driver's SA under this perspective.

The computing perspective argues that rather than examining the levels of SA held by individual road users, it is more appropriate to study and enhance the SA presented by different displays, in-vehicle technologies and road infrastructure (Salmon et al., 2009). Relating this approach to Endsley's model, it is suggested that directly presenting the higher levels of SA (Level 2 and 3; SA-related information) on displays augments SA acquisition and maintenance since it removes the need to interpret Level 1 data. For instance, Endsley & Jones (2011) proposed that information presentation can support SA. Endsley & Jones (1997) reported that directly providing the time and distance to travel on the basis of available fuel is more useful than informing the pilots of the fuel levels and speed. This form of presentation of information facilitates Endsley's Level 2 and Level 3 SA, as need for Level 1 is eliminated by presenting the precise information needed. These results have been translated to in-vehicle systems like dashboards of cars

and buses. Thus, engineers or operators usually "set their SA" in the instruments there to assist their operation. This leads them to spend less effort in visually searching for any object in the road environment which might be out of ordinary. Thus, under this school, the driver totally relies on technology to maintain their SA. Clearly, this leads to some very profound problems with automation widely discussed in human factors literature. The operators start trusting the settings of the instrument more than themselves, instead of looking for cues themselves which are worth being alarmed (Sorensen, 2011).

Thus this perspective is technology-led, and focuses on networks and information. SA is achieved by integrating countless technologies to provide operators with access to information based on their circumstance (ESRI, 2008). SA as a technology-led phenomenon tends to focus on the role of geographic information systems which are concerned with representation of terrain, maps, and infrastructure and so on (Ackerman, 2000). It is claimed that these technologies tell people where they are and where their resources are located, which maintains the operator's SA (Carroll, 1999). As ergonomists we very often have to deal with aviation literature which is also influenced by inputs from the computing perspective. Engineers in aviation are frequently reported to have presented SA technologies (DeMeis, 1997) as if the technologies are containing the SA. For example, Ground Proximity Warning Systems (GPWS) have been developed to prevent controlled flight into terrain (DeMeis, 1997).

Endsley & Jones (1997) suggest that the manner in which information is presented in such accessory systems influences SA by determining how much information can be acquired, the accuracy of the acquired information, and the degree of compatibility of acquired SA information with SA needs. Furthermore, Albert et al. (1999) recognise the timeliness, accuracy and relevancy as dimensions of the information presented. This means merely increasing the information presented won't guarantee better decisions (Bolia et al., 2006) On the contrary a burst of information to the operator may lead to information overload, poor interpretation, and the presence of non-relevant information that is more likely to degrade rather than enhance SA. So, the concept of "the more the better" is not applicable under this perspective.

The similarity with Endsley's model can be seen when it comes to design. The traditional views of SA given by Endsley (1999) aims to design systems or assistance in a way that it can contain the shortcomings of human beings or operators who are driving them. Again the difference of the computing/engineering perspective when compared to the psychological perspective can be seen in terms of principles of design. As the psychological perspective is based on the idea of SA being born out of cognitive processes within the individual, the design aim is to improve performance by advanced training of the operators. On the other hand, the computing perspective which believes that SA is not within the individuals but in the artefacts contends that SA can be improved by technological advancements of the systems only. The latter can be readily seen in everyday life, and also complements the layman understands of ergonomics.

In contrast to the two perspectives presented above, the following section presents an approach which does not separate individuals from the system, but views them as a part of the system. The subsequent section clarifies that SA is neither within the individual nor with the system but is born out of interaction between the two.

2.3.3. Interactionist/Systems/Situation Perspective

Psychological and engineering perspectives provide very interesting viewpoints and insights to situation awareness; however, they are so contrasting in nature that following only one in a real-time environment won't lead us to eliminate deaths due to road accidents. Thus, the insights from the above two schools call for one intermediate school or perspective, which can act as a link between the two and at the same time is practical enough to apply in real-time environment. Thus, the following defines the situation perspective to situation awareness:

"*Complex systems cannot be understood by studying parts in isolation. The very essence of the system lies in the interaction between parts and the overall behaviour that emerges from the interactions. The systems must be analysed as a whole*" (Ottino, 2003, p.293).

Hollangel (2001) explained that people and artefacts comprising a system together form a joint cognitive system, and the cognitive processes that arise from the system are

distributed across the system (Salmon, 2008). This implies that cognition in a system is not maintained in isolation but through coordination between the system units (Stanton et al., 2006). The system units are comprised of both human and non-human agents (Salmon et al., 2009; Salmon et al., 2009; Stanton et al., 2010; Salmon et al. 2011) like displays or artefacts particularly used by the engineering perspective to maintain SA.

Despite the name 'systems ergonomics', this does not discount the role of individuals in maintaining SA. The difference here is that the individual in systems ergonomics is merely considered as part of the system like other agents (traffic signals, roundabouts, safety margins, etc.) rather than being the centre of attention. This is a major change from the psychological perspective, and explains the role of the individual with the help of schema theory. According to the schema theory, the individual hold schemata which are activated by the task which is being performed (Stanton et al., 2011). Through task performance the phenotype schemata are created through the interaction between people, world and artefacts (Salmon et al., 2009).

It is quite clear that the systems/situation perspective adopts a very holistic approach to studying SA in systems (like roads, railways, nuclear power plants, control rooms and many more). Following the systems perspective there are various technical and social agents (like human road users) in a system that work and coordinate their SA effectively as a team. Due to appearing chronologically before others, most of the human factors literature on team operations has been theoretically underpinned by Endsley's three-level model. So, Endsley's model which was originally meant to understand SA of the operator was used to understand SA of teams.

Before a system's perspective team SA was poorly understood by human factor scientists and practitioners. Applying individualistic models (e.g. the three-level model) of SA to understand SA in teams will not yield accurate results. Team SA is far more complicated than individual SA; hence the psychological perspective is not sufficient to explain team SA. Team SA was initially considered to be the summation of SA of individual team members' SA. Salas et al. (1995) argue that team SA is far more complex than this addition of individual SA. With the exception of Salas et al. (1995) various models to explain team SA report team SA to be a common picture between the

members (Wellens, 2003; Nofi, 2000; Stout et al., 1994). The upgraded version of those models is the team SA model given by Salas et al. (1995).

Team SA according to this model depends on communication between team members at various levels. Perception of SA elements is influenced by goals and tasks of each member in a team. This perception is also influenced by team performance factors and team capability (Stanton et al., 2006). The comprehension of this information is influenced by the interpretations of other team members. Strangely enough, this means that "SA leads to SA and modifies SA" (Salmon et al., 2009, p. 21).

The major team SA models mentioned later in section 2.6 are designed to measure individual SA. Individual SA is developed and shared with team members, which then develops and modifies team members' SA (Salmon et al., 2009). In contrast to this, situation awareness transcends boundaries of operators of a system and becomes a utility that is achieved and maintained by coordination of human and technological agents working as a team within a system. If the coordination is hindered by some factor then the entire system's functioning comes to a stop or is affected. This also forms the crux of the systems perspective and this thesis. Finally, the models given before Salas's model of team SA is based on shared understanding of the same situation. For example, Perla et al., (2000, p.17) suggests that "we all understand a given situation in the same way." According to Stanton et al. (2006) no set of experiences is likely to be identical, and thus this follows for each of the interpretations. Therefore, team SA models (Wellens, 2003; Perla et al., 2000; Stout et al., 1994; Salas et al., 1995) heavily rely on the psychological perspective of SA; none of the team SA models can be categorized under systems perspective, however all of them purport to measure team SA. This idea of measuring SA of team corresponds with the system/situation's perspective's aim of measuring SA of the team but still cannot be considered under system's perspective.

The distributed situation awareness (DSA) model given by Salmon et al. (2009) provides an alternative way of assessing team SA. The main idea behind this model which blends it perfectly with the systems perspective is 'compatibility' between social and technical agents. DSA does not only explain how people work in teams but how SA

bonds humans and technology together (Stanton et al., 2006; Stanton et al. 2009). For example, in a road environment, there is an exchange of information between drivers, pedestrians, traffic lights, road materials and other infrastructural elements. In this way a road is considered as a whole system with different agents (like signage, pedestrians, lights and other similar design elements) in it and no agent is given importance over others.

According to Stanton et al. (2006, p.1291), distributed SA can be defined as "activated knowledge of a specific task at a specific time within a system". This means that information held by the system becomes active at different points in time based on the goals and activities being performed (Salmon et al., 2009). Each individual holds different SA for the same situation, depending on his or her activities or goals (Salmon et al., 2009). Thus, SA is goal-dependent. If there is an incompatibility between these goals, the functionality of the system will break down and there can be an accident. It is interesting to note that, following the same line of goals, Banks and Millward (2009) argue that mental models need not be contained within an individual, and that there is a possibility of these being distributed in a group (i.e. shared mental models). Each human agent holds a part of that mental model (Bank & Millward, 2009). As mentioned above, the parts of shared mental models held by individuals can be used for communication and interaction where necessary, and this can also be a form of SA transaction.

In conclusion, the concept of distributed SA is the latest model of explaining and developing SA in joint collaborative systems such as roads. The systems perspective is certainly ahead of the ever-popular psychological perspective, and gives a complete picture of the environment. The Systems/Human Factors school of thought transcends the fragmented or incomplete picture painted by either the engineering or individual approaches, by providing means to view the system as a whole (Stanton et al., 2010). It considers the information held by artefacts and humans fairly and explores the way they interact. Therefore, systems perspective and the DSA model will be used in this thesis to understand interaction amongst road elements and SA.

2.4 Definitions of Situation Awareness

Defining SA is no simple task because the term means many things to many people. Any definition general enough to account for anything and everything related to driving a car or designing driving aids might prove too broad to be useful (Dominquez, 2004). On the other hand, overly specific definitions often prevent us from transferring concepts, issues, and results from one domain to another (Dominquez, 2004).

From its inception, situation awareness has been a debatable concept. There have been numerous attempts (e.g. Adams et al., 1995; Bedny & Meister, 1995; Dominquez, 1994; Gugerty, 1997; and many more) but none of these definitions can be referred to as a universal definition of situation awareness. Most of the researchers and human factors practitioners refer to Endsley's definition of SA. Additionally, only Endsley's definition of SA has received 50 citations in peer reviewed journals relating to ergonomics research. Despite such popularity Endsley's definition is not a universal one.

There are many factors which have hindered in the development of a universal definition of situation awareness. One of them is the similarity of SA with other concepts. For instance, the term situational assessment is often confused with situation awareness. Situational assessment (also abbreviated as SA) is referred to as a process for achieving situation awareness (Endsley, 1995). These processes or situation assessments of developing and maintaining situation awareness vary widely between individuals and contexts (Endsley, 1995). Situation awareness is also confused with mental models (Sarter & Woods, 1991) and sense-making (Salmon, 2008). The latter is referred to as a "motivated continuous effort to understand connections in order to anticipate their trajectories and act effectively" (Klein et al., 2006, p.71). Endsley (2004) pointed out that sense-making is a subset of the processes to maintain situation awareness and not situation awareness itself. Furthermore, Endsley (2004) reported SA as forward-looking, projecting what is going to happen in order to inform effective decision making processes, but sense-making to be retrospective process focusing on reasoning of past events. Due to the similarity of these concepts SA means different things to different people, which has created a barrier for adopting a universal definition.

The product versus process debate on situation awareness has added to the already existing controversies around the area of situation awareness which has also contributed to not having a universal definition of SA. 'Process' refers to the cognitive activities involved in revising and attaining the state of SA, whereas 'product' refers to the state of SA with regard to the information and knowledge (Stanton et al., 2001). The vast majority of research has concentrated on measuring SA as a product, triggered by Endsley's research. However, Andre (1988) argues that product research has avoided the issue of how information was obtained, or resources or workload constraints on processing that information. Banbury et al. (2001) has also reported that in order to develop the 'product', mental dynamism and constraints need to be understood. For instance, to understand how a driver's SA is formed it is essential to consider the design of the vehicle, which elements are more important to driver than the others, what are his expectancies from the road, and so on.

The psychological perspective is based on product S.A; however there is disagreement amongst some definitions on SA being the product of a process. For example, contrary to the typical product SA definition, Sarter and Woods (1991) propose a classical process-oriented definition of SA. Sarter and Woods define SA as "the accessibility of a comprehensive and coherent situation representation which is continuously being updated in accordance with the results of recurrent situation assessments" (1991, p.52). The psychological perspective often provides a rich description of elements of decision-making activities in complex systems such as perception, comprehension and projection as suggested by Endsley. Following the process approach, Endsley contradicted herself stated that, "situation awareness can be described as person's state of knowledge or mental model of the situation around them" Endsley (1995, p.18). Isaac (1997, p.185) further added to this confusion by defining SA as a product as well as process, stating that:

> *SA refers to a cognitive state or process associated with the assessment of multiple cues in a dynamic situation. It may refer to a person's knowledge and reference to their status within a space and time continuum (pilot) or an operator prediction within a known space and time continuum (air traffic controller).*

Although it is the most highly cited definition, many of these citations are to point out the limitations in the approach. The different definitions of SA indicate the incoherent nature of SA in the psychological school of thought. In contrast to the psychological perspective, the situation perspective has no incongruence in definitions and refers to SA as a product as well as a process. It offers an explanation of cognitive activity used to attain SA and details of what SA as a product contains. Smith & Hancock (1995, p.142) define SA as "externally adaptive consciousness that has its product knowledge about a dynamic task environment and directed action within the environment." Adams et al. (1995) modified Neisser's perceptual cycle and reported SA to be a product as well as a process. Adams et al. (Uhlarik, 2002, p.4) propose that "SA as a product is a state of currently activated schema, and as a process SA is the current state of the entire perceptual cycle". For emergency conditions, Adams et al. gave a much more elaborate model to adequately capture behaviour.

Distributed Situation Awareness defines SA as "activated knowledge of a specific task at a specific time" (Stanton et al., 2006, p.1291). This definition means that information required for SA becomes active at different points in time based on goals and activities being performed and their requirements (Salmon et al., 2009). Again DSA is referring to SA as a product and a process. This is similar to Adams et al., and Smith and Hancock's idea of SA being product and process. Moreover, the definition of SA provided by Stanton et al. (2006) is a refined form of the description provided by Cowan (1988) for activated working memory as awareness. The definition of SA given by Stanton et al. (2006) does not only fit in with systems perspective but is also parallel with the objectives of this thesis. Therefore this thesis will follow a systems/situation perspective, by adopting a DSA model and using the definition of SA developed by Stanton et al. (2006). For the purpose of this review over 30 definitions of situation awareness were identified. Definitions were gathered from books, journal articles, conference proceedings, magazine articles and technical reports. It is beyond the scope of this PhD to discuss each and every definition of situation awareness, however definitions which occur often have been critically analysed. Two tables of SA definitions are presented: one with individual SA (2.2) definitions and the other with team SA definitions (2.3).

Table 2.2. Summary Table for Definitions for Individual or General Situation Awareness (Salmon et al., 2009)

Definition	Author	Theoretical Underpinning	Origin	Citation[1]
Perception of elements, comprehension of meaning and projection of future status	M.R. Endsley (1995)	Three-level model	Started with aviation, later spread to other domains	2253
Situation awareness is adaptive, externally-directed consciousness that has its products knowledge about a dynamic task environment and directed action within that environment.	Smith & Hancock (1995)	Perceptual cycle model	Air traffic control	204
Situation awareness is based on the integration of knowledge resulting from recurrent situation assessments.	Sarter & Woods (1991)	Working memory, mental models, situation assessment awareness	Aviation	273
"…an abstraction that exists within our minds, describing phenomena that we observe in humans performing work in a rich and usually dynamic environment."	Billings (1995)	Information processing theory	Aviation	31
Situation awareness is knowledge of current and near term disposition of both enemy and friendly forces within a volume of airspace.	Hamilton (1987)	Three-level model	Military airspace	6
SA contributed to good performance, it is not synonymous with it. It is possible to have good SA but still not be a good pilot because of poor motor skills, coordination or attitude problems.	Adams, Tenney & Pew (1995)	Perceptual cycle model, working memory	Aviation	33
SA is the pre-requisite state of knowledge for making adaptive decisions in situations involving uncertainty, i.e. a veridical model of reality.	Taylor (1990)	Theories of attention and cognition	Military, air traffic control and nuclear power	272
Activated knowledge, regarding road user tasks, at a specific time, within the road transport system. From a road user perspective, this knowledge encompasses the relationships between road user goals and behaviour, vehicles, the road environment and infrastructure.	Salmon, Stanton & Young, 2012	Distributed Situation Awareness	Road Transportation	13

Table 2.3. Summary Table of Definitions for Team Situation Awareness Source: (Salmon et al., 2009)

Theory	Authors	Theoretical Underpinning	Origin	Applications
Team SA	Endsley & Robertson,(2000)	Three- level model	Aviation	Military aviation maintenance
Inter- and intra-team SA	Endsley & Jones, (2001)	Three- level model	Military	Military
Team SA	Salas et al. (1995)	Three- level model, Team work theory	Generic	None
Team SA	Wellens, (1983)	Three -level model, distributed decision -making model	Military	Military
Distributed cognition approach	Artman & Garbis, (1998)	Distributed cognition theory	Tele-operation	Tele-operations
Military awareness team SA model	Shu & Furuta, (2005)	Three level model , shared cooperative activity theory (Bratman, 1992)	Process control, Artificial intelligence	Process control
Distributed situation awareness model	Salmon et. al., (2009)	Perceptual cycle, Distributed cognition theory (Hutchins, 1995)Distributed SA theory (Artman & Garbis, 1995)	Maritime	Military, maritime, energy distribution, aviation, air traffic control, emergency services, driving

[1]**Citations from Google**

2.5 Theoretical Underpinning

After reviewing the three contrasting perspectives of situation awareness, it can be concluded that situation awareness is a multi-perspective area of research. No one definition agrees with the other. It is interesting to note that domains which have found SA to be critical essentially involve teams and not individuals (Stanton et al., 2008). This means, if individual perspectives are applied to study behaviour in teams and team SA, the shortcomings of individual perspectives will very quickly become visible (Stanton et al., 2008). A long legacy of SA in aviation has thus played a pivotal role in manifesting Endsley's model for explaining SA to researchers. This lack of consensus for a unified approach to SA also exists for theoretical underpinnings, as each perspective to SA follows a unique underpinning or a combination of underpinnings. For instance, Endsley's model (and similar models covered above) stemmed out from mental models and bonded with information processing. Similarly, team SA models are based on mental models and shared mental models. However, distributed situation awareness (Salmon et al., 2009) has the foundation of Hutchin's (1995) distributed cognition, schema theory and Neisser's perceptual cycle, which are actually linked to each other.

The following section will explore these four theoretical underpinnings, i.e. mental models, information processing, schemas, and the perceptual cycle used by different researchers to make their models theoretically robust.

2.5.1 Mental Models

Although the term 'mental models' is pervasive in cognitive psychology and human factors literature, there are surprisingly few explicit definitions for this (Rouse & Morris, 1986). This lack of definitions is an indicator of the acceptance of the construct on an intuitive basis (Rouse & Morris, 1986). Nevertheless, it is interesting to consider the few definitions for mental models which appear frequently in literature of various areas of research.

"The mental image of the world around us that we carry in our head is a model. One does not have a city, a government or a country in his head. He has selected concepts and relationships, which he uses to represent the real system" (Forrester, 1971, p.213). Forrester (1994) later criticised mental models due to the immense difficulty it poses while measuring those concepts and relationships between various variables he mentioned before. This limitation has been agreed upon by numerous scientists (Dorner, 1980; Sterman, 1989; Brehmer, 1992; Smith & Hancock, 1995). Meadows et al., referred to mental models as "intuitive generalizations from observations of real-world events" (1974, p.4). According to Rouse and Morris (1986) mental models are "mechanisms whereby human beings are able to generate descriptions of system purpose and form an explanation of system functioning and observed system states and thereby predict future system states" (p.351). Senge (1990, p.6) defined mental models as "deeply imagined generalizations and assumptions or pictures or images that influence how we understand the world and how we take action. Very often, we are not consciously aware of our mental models or the effects they have on our behaviour". The common theme in all the definitions mentioned above are describing, explaining and predicting (Rousse & Morris, 1986). Refer to Figure 2.6. The operators form a conceptual equivalent of the external world to predict system behaviour.
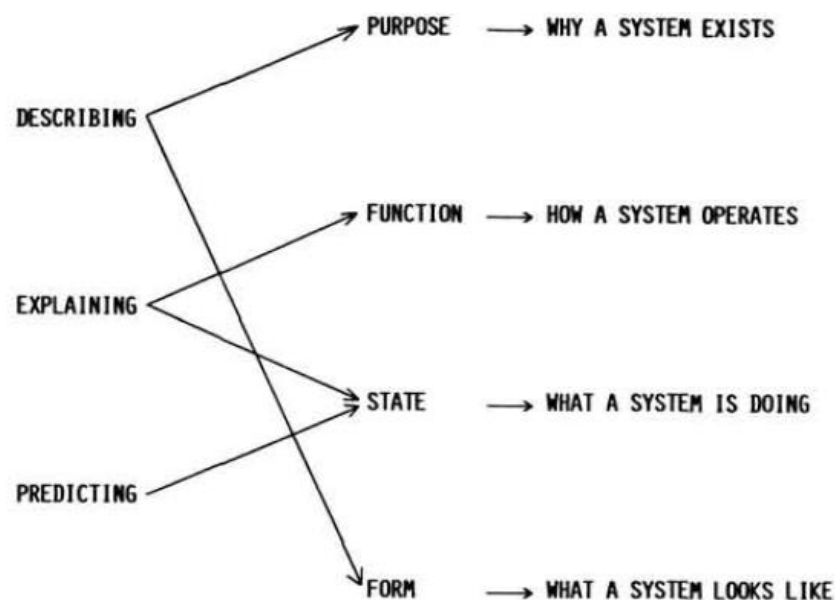
Figure 2.6. Similarity in Definition of Mental Models (Rouse & Morris, 1986, p.351)

This common theme of prediction fits very well with Endsley's three-level model of SA. Endsley's definition of SA as the "*perception* of elements in the environment within a volume of time and space, the *comprehension* of their meaning and *projection* of their status in the near future" (1995, p. 35) has an uncanny resemblance to Rouse and Morris's (1986, p.351) definition ("mechanisms whereby human beings are able to generate *descriptions* of system purpose and form *explanation* of system functioning and observed system states and thereby *predict* future system states"). This indicates that there might be an overlap between mental models and SA. Moreover, Sarter and Woods (1991) have discussed the role of various constructs in SA formation and have cautioned researchers to differentiate SA from mental models.

Mental models are used as experiences to facilitate the acquisition of Level 2 data in Endsley's model. Mental models are very critical to explain the development and maintenance of SA in accordance with the three-level model. According to Endsley, features of the environment are mapped to mental models in the operator's mind. These models are used to facilitate the development of SA (Salmon et al., 2009). Endsley suggests that more experienced operators use mental models to facilitate the integration of information of Level 1 SA to Level 2 SA (Salmon et al., 2009). This implies that mental models formed by training and experience help an operator to direct attention to critical elements in the environment shown in Figure 2.6 (why a system exists, what a system does), integrating the information to aid understanding of the environment (how a system operates, what a system is doing) and generating possible future state and events (what a system can do or is most like to be doing).

In aviation literature mental models are referred to as a picture of the real situation in the minds of the pilots and air traffic controllers. The loss of that picture results in loss of SA. Smith and Hancock (1995) have argued that this picture in the pilot's mind is not SA, but it is SA which helps the pilot to develop this picture. Applying the concept of mental models in the context of driving cars will mean that drivers should be having a picture of the road environment in their minds. This mental picture is created in their head with the help of information presented to them in the form of fuel level, speed, global positioning system, etc. The information generated by these gadgets will help the driver not only to maintain a mental picture but to help in conceptualisation and

prediction of traffic movements. Thus, this mental picture is created methodically, by contrasting the possible course of action with the other available alternatives using a common set of abstract evaluation dimensions.

Recognition-primed decision making (RPD as given by Klein et al. (2004)) disagrees with this methodical approach of selection of an appropriate mental model from a library of models inside the operator. RPD describes how an operator chooses the most appropriate course of action for a situation rather than comparing and contrasting mental models to arrive at an action. Immediate selection of an action will depend on an operator's experience and training. This correlates with Logan's (1988) view of the cognitive processing of experts. Logan (1988) reported that the cognitive processing of operators at an expert level is automatic through a single step, from memory to action. It is worthy to note both RPD and mental models highlight the role of long term memory and experience while explaining the difference in performance of experts and novices.

One mental model is often not sufficient to explain SA formation in teams (Salas et al., 1995). In such a circumstance mental models are shared between team members. Shared mental models refer to the degree to which each person shares a similar representation of the team with other team members. Shared mental models have also been reported to be one of the many factors contributing to high SA (Ensley & Bolsted, 1999; Endsley & Jones, 1997). Mental models have been regarded as the theoretical underpinning for both individual and computational perspectives. In other words, it will be fair to say that mental models form a theoretical underpinning for models from Endsley and Salas et al. to explain situation awareness in complex systems.

Hence, mental models can only be considered as a theoretical underpinning to explain SA formation for the Endsley and Endsley–type models. It cannot be considered as the theoretical underpinning of SA formation in general, or for other models and perspectives advocating for SA in complex systems. Also, the mental model is not a well-understood factor. Smith and Hancock, 1995 referred to it as ill-defined and problematic; using it as a theoretical underpinning may lead to explanation of higher-order processing to explain SA.

2.5.2. Schemas

As discussed by Stanton et al. (2008), the early mention of schema theory is found in Head (1920) and Piaget (1926). Piaget (1950, p.201), reported "schema to be basic building blocks of knowledge and development". According to Bartlett (1932), schema is the active organization of past reactions and past experiences which are combined with information in the world to produce behaviour. Schmidt (1975) argue schema to be a set of rules or abstractions to determine a movement. A more colloquial definition for schema might be that "schemas are slogans or picture we make up in our mind about everyday situations as we experience them again and again" James, 1997 The literature on schemata basically divides schemata into two categories. The first category which consists of early definitions describes schemata as a 'general category of mental representations' (Evans, 1967; Rumelhart & Ortony, 1977). The second more updated category includes the concept of knowledge stores to explain abstract ideas like procedural knowledge, which dictate how to do things rather than just stores of discrete categories of things (Brewer & Nakamura, 1984, Schmidt, 1975). Thus, all of these different definitions of schema summarize that it is an organisation of knowledge in a convenient and easily comprehensible manner which directs an action in a particular situation. This means that any error in organisation and selecting scripts (functional units of schema) may lead to erroneous action.

When a person carries out a task, schemata affects and directs how they perceive information in the world, how this information is stored, and how it is activated to provide them with past experiences and the knowledge about the actions required for a specific task (Mandler, 1984). This means that schema helps us to process information quickly and efficiently to facilitate memory recall. Thus it is possible to remember details that are consistent with our schema, more easily than for those that are inconsistent (Hastie & Park, 1986; Stangor & McMillan, 1992). Van Viler and Schermers (2000) underlined homogeneity and familiarity as essential features of SER. As such, SER attempts to make the road environment much more consistent with the existing schema. Self-explaining roads are in line with the expectations of road users (Theeuwes & Godthelp, 1995). First applied in the Netherlands, SER relies on changing the visual characteristics of the road to influence the behaviour of road users (Charlton et al., 2010).

The role of schemata in naturalistic settings where decisions are to be made in real time is best explained by Klein et al. (1998). In this explanation, the cognitive world is organized with concepts, ideas, and cognitive constructs such that they narrate a story to its user. This means that schemata are influenced by how information is presented in the environment. This idea of environment-sensitive schema has been exploited in Neisser's perceptual cycle. Schmidt (1975) outlined initial conditions, parameters or variant features, sensory consequences, and response outcomes as requisites for forming a schema. Schema is seen as a form of mental structure (Brewer & Nakamura, 1984; Rumelhart & Ortony, 1977) whereas Evans (1967, p.87) argues schema to be less of structural entity in the mind and more as "rules serving instructions".

Although there is a substantial argument regarding the nature of schemata, the characteristics of schemata are generally agreed upon (Plant & Stanton, 2011). According to Anderson (1977), schema has five defining characteristic features. They are meaningfully organized, embedded with other schemata, and contain sub-schema; they are dynamic as they change from moment to moment as information is received; they are restructured when incoming data needs to be reorganized; and finally they are based on gestalt mental representations. Other theories on schema (Norman, 1981; Rumelhart, 1984) also emphasise similar characteristics, though the terminology might differ (Plant & Stanton, 2011).

The concept of schema has always been around. According to Marshall (1995) the early mentions of schemata are traceable from the writings of Aristotle, Plato and Kant. Barlett (1932) brought schemata into focus. This is very similar to the concept of SA itself, which has always been around but was brought into focus by Endsley (1995). The contemporary literature of schema by Neisser (1967; 1976) established schemata as an important construct in cognitive psychology. There have been numerous studies which have empirically tested schema (Edmonds & Evans, 1966; Posner & Keele, 1968; Posner & Keele, 1970), and this is analogous to the evolution of SA. Schema, now a well-established concept, has many branches. Motor schema theory (Schmidt, 1975), gender schema theory (Bem, 1981), schema theory in clinical practice (Young et al., 2003), driving (Hole, 2007), tool use (Baber, 2003), and politics (Axelrod, 1973).

Schema provides very interesting insights into the role of past experiences in guiding cognitive processes and modifying messages from environment. The notion of schema or mental representation is well established, even though there is still debate as to how these representations are developed and maintained (Woods et al., 2010).

2.5.3 Difference between mental models and schemas

According to Cheng et al. (1986), schemas are abstract knowledge structures derived from ordinary life experiences. They can be applied across several domains but cannot be universally valid like inferences or rules (Halford, 1993). Johnson-Laird's mental model (1983) is very similar to this. Schemas and mental models produce representations that have a very specific semantic content. However, according to Johnson-Laird, mental models are constructed so as to represent the possible ways to combine premises. A new mental model is constructed each time a new problem is encountered (Halford, 1993), whereas in the case of schemas an existing schema is introduced from previous experience to represent the problem. Hence, mental models are always constructed to fit specific tasks and schemas are taken from past experience.

Mental models and schemas are both representations, but mental models are at work while an operator is working on a task. On the other hand, schemas structure those representations of a problem from the knowledge available from experience. Thus, schemas provide a template for going about a task rather than a detailed description of task performance.

Mental models are internal representations (Vega et al., 1996) of the external world. They exist in the minds of the individuals. This property of mental models makes it difficult to become a subject of direct introspection and measurement (Jones et al., 2011). Schemas on the other hand being widely researched have more documented measurement tools. For example, Wicks (1986; 1992) has devised a method to assess schemas in mass communication media. Moreover, schemas are often networked and measured.

Schemata adapt to new information much more readily than mental models. Mental models are resistant to change, while a new schema can be created and modified readily. Finally, schema provides "human in the system" explanation for any error (Plant & Stanton, 2011) whereas mental models provide "system in the human" explanation for any error. This implies mental models are suited for the explanation of Endsley and Endsley-type models, whereas schema can give an explanation for a more generic idea of situation awareness.

Finally, there is account of shared mental models, but there is no account of shared schema (Mohammed, 2001).

2.5.4. Perceptual Cycle
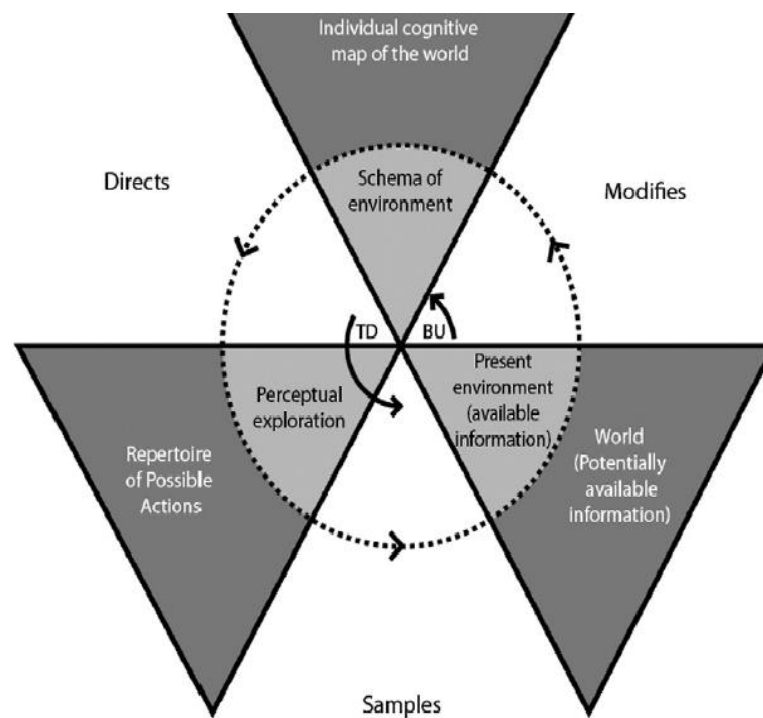
Figure 2.7. The Perceptual Cycle Model (Hancock & Smith, 1995)

The perceptual cycle model as seen in Figure 2.7 is the most cited work on schema (Stanton et al., 2008). First cited in *Cognition and Reality* by Ulrich Neisser in 1976, it has been widely applied ever since. The most striking aspect of this model is the adoption an ecological approach which is in contrast with the information approach adopted by models of individuals' perspectives like Endsley's three-level model. The

ecological approach suggests perception to be an active rather than passive process (Neisser, 1976; Hayes, 2005; Salmon et al., 2009; Plant & Stanton, 2011; Stanton et al., 2010). Perception is regarded to be a guided exploration following the direction of active schemata (Salmon et al., 2009). This kind of exploration leads to the adaptation of the environment by the perceiver which guides future exploration. In the perceptual cycle model, the definition of schemata was updated to explain "schemata as active structures that guide the exploration and interpretation of the information which in turn changes those structures, further guiding exploration and so on" (Salmon et al. 2008, p.222).

Since schema is so critical to the perceptual cycle model, the form and nature of schema have the ability to determine what we are able to perceive, i.e. how it fits to our own personal schemata through the interactions between people and environment. Neisser (1976) identified and distinguished between genotype and phenotype schema (Stanton et al., 2009). The precursor to genotype schema is the global prototypical routine (GPR), while the precursor to phenotype schema is the local state specific routine (LSSR). GPR might be thought of as the schemata in the mind of the person whereas LSSR can be referred to as an activated schemata as applied to a specific problem by a user (Baber & Stanton, 2002). Baber & Stanton (2002) suggested that the interaction between the two leads to the modification of the schemata towards the goal.

Genotype schemata analogous to GPR refer to the systematic factors that are directly proportional to the development of an individual's cognitive phenomena and behaviour. A more formal definition given by Plant & Stanton (2011) defines genotype schema as residual structure in the mind that directs our activity in the world. Genotype schemas are the templates for action and have the possibility to develop, but the determinant of their development is in the interaction of those templates in the environment. Stanton et al. (2010) argue that that one of the key characteristics of socio-technical systems is the emergence of behaviour from such interactions, and these interactions need to be understood.

The phenotype schemata are analogous to LSSR and mental models. It is the "in the moment" schema and comes about as action schemata in the world. This is where the role of expectation and experience influences our active exploration of the world. The

initial triggering of the schema is a bottom-up (BU) process produced from situations within the environment. These initiate schemata that are based on past experiences and expectations at which point this becomes a top-down (TD) process.

In the perceptual cycle model the main cognitive structure that determines process, movement, and action is called anticipatory schemata (Chimir et al., 2005). The reciprocal cyclical nature between person and environment forms the basis of the perceptual cycle model, as well as the distributed situation awareness model of the systems perspective. Neisser gave his model 32 years before it was considered as a robust theory to pin down DSA (Salmon et al., 2009). Despite this, ideas related to the perceptual cycle model have always been used for accident analysis in safety-critical domains (Hall & Silva, 2008) like rail (Stanton & Walker, 2011) and driving (Bellet et al., 2009).

Neisser (1976) states that perception is an active process directed by awareness. It is active because it emphasises perceiving some objects in more detailed ways than others. Neisser (1976) also points out that breaking the cycle into individual parts will lead to loss of perception. According to the perceptual cycle, schema supports the observer in perceiving the world through directing, for example, what to see. In the light of a driving scenario, if a driver has switched from a motorway to a country road, there will also be a change in schema. In the case of the country road, the driver would expect to see fewer cars and narrower roads in comparison to the motorway. This is due to the fact that the driver has already learnt these facts and figures about various roads of the rural road environment where he/she is driving. This does not mean that the driver would only see what he/she is expected to see. For instance, at the Kerang railway line accident the truck driver failed to respond to the crossing warning device due to activation of incorrect schemata (Salmon et al., 2012). The schema for the rail level crossing failed to activate timely. That is, the truck driver perceived the crossing as 'no train approaching', which, through shaping the truck driver's expectations and perception, then caused a 'look but failed to see' error in which the driver scanned the rail level crossing safety warning lights, but did not perceive their flashing state (Salmon et al., 2012). This is called looked but failed to see (LFBTS) errors as given by Staughton & Storie, 1977.

As illustrated in Figure 2.7 above, Neisser (1976) presented the view that the humans are thought to be closely coupled with a person's interaction with the world, a concept fundamental to system's approach. Both (human and the environment) inform each other in a reciprocal and cyclical relationship. World knowledge' that is schemata leads to anticipation of certain types of information and provides a way of interpreting that information (anticipatory schema) (Stanton et al., 2009). The environmental experience results in the modification and updating of cognitive schemata and this in turn influences further interaction with the environment (Stanton & Plant, 2011).

A major strength of the perceptual cycle model is that it provides a human in the system approach, whereby the schemata are influenced by the environment and can constantly react to changing environmental conditions. The dynamic cyclical view of person and world interaction presented by the perceptual cycle model opposes the traditional view of Endsley's model wherein problems reside in the humans, or according to the computational perspective problems resides in the system. Instead, the perceptual cycle model accounts for the interactions that occur.

Like Endsley's three-level model which forms the foundation for the psychological perspective, Neisser's perceptual cycle is the foundation for systems perspective. It links, "mind to action to world to back to mind" (Stanton et al., 2010, p.34). There are considerable differences between the perceptual cycle model and the foundations of the psychological perspective. Firstly, the perceptual cycle is a combination of bottom-up and top-down processes whereas the three-level model is a bottom-up process only. Bottom-up processes are data-driven, and top-down processes are goal-directed (Endsley & Garland, 2000). Endsley's model claims to be an alternation of top-down and bottom-up processes, and this alternation is said to be critical for achieving SA (Endsley & Garland, 2000; Endsley, 1988; Endsley, 1995). Endsley (2000) claims that the perception of elements (Level 1) is goal-driven; goals also simultaneously act as a filter for the comprehension of perceived information. It dynamically alternates to data-driven processing when the perceived elements act as new goals to be achieved. Theoretically it is true and attractive, but in reality, Endsley's model does not seem to display this complexity. Very often drivers are simply aware of objects without conscious processing (Young & Grenier, 2009) which affect their driving behaviour.

A significant aspect of human factors literature is that relating to affordances (Gibson, 1977). Affordances are directly perceived qualities of a system, and are dynamic properties that can influence the potential states of a system in future (Young & Grenier, 2009). The state that will be realised will be complementary to the intentions of the individual (Stoffregen, 2003). Thus Endsley's model can be regarded as a bottom-up process only, and not an alternation of both. Also, due to this limitation and the unusual nature of affordances, Endsley's model cannot be applied in interface design (Young & Grenier, 2009). Neisser's perceptual cycle on the other hand explains SA to be an integration of top-down and bottom-up processes. Initial triggering of the schema is a bottom-up process produced from a situation within the environment (Stanton & Plant, 2011), and these initiate schema that are based on past experiences and expectations at which point the process becomes top-down. This modification of schema makes it resistant to the dynamicity which affordances have to offer.

2.5.5 Distributed Cognition

Rogers (1997, p.1) defined distributed cognition as a "hybrid approach studying cognition at social, organizational and cognitive perspective". Rogers (1997) also claimed the distributed cognition approach to be the most widely-known level of analysis for studying "complex socially distributed cognitive activities" (Rogers, 1997, p.1). Distributed cognition forms the foundation for the distributed situation awareness model (Salmon et al., 2009) which is adapted on within this thesis. However, Artman and Garbis (1998) applied distributed cognition to study teams for the very first time. Distributed cognition principles can also be seen in systems/situation perspective.

The distributed cognition approach is quite contradictory to all other cognitive science approaches in the 1980s and even prior to that. Like all other approaches in cognitive science, distributed cognition also focusses on representing cognitive processes (Hutchins, 1995). The traditional view of cognition as localized phenomenon best explained in terms of information processing by the individual was challenged by Hutchins (1985) through distributed cognition (Rogers, 1997). The theoretical foundations of Hutchins' distributed cognition are based on the cognitive sciences, cognitive anthropology, and social sciences (Roger, 1997).

What sets distributed cognition apart from other cognitive science approaches is its broader view of cognitive events (Hutchins, 2000). Mainstream cognitive science looks for cognitive events in the manipulation of symbols (Newel et al., 1989) inside the individual actors. In other words, "distributed cognition does not expect all such events to be encompassed by the skin/skull of individual" (Hutchins, 2000, p. 1). According to Hutchins (1995), when this fundamental principle is applied in cognition in the wild three distinct cognitive processes become apparent:

1) Cognitive processes may be distributed across the members of a social group.
2) A distribution of cognitive processes can be seen as an interaction of external and internal events.
3) There is also a possibility of transformation of products of earlier events to shape future events.

Hutchins (1995) reported these three aspects are essential to understand human cognition. To apply a distributed cognition approach, a system is supposed to be a joint cognitive system. A true cognitive system should possess three properties to be termed as such. Rogers (1997) underlined these properties to be the inclusion of more than 2 people, with knowledge possessed by the members of the cognitive systems being highly variable, and finally that the system will provide access to information to all members of the system. Individuals working together on a collaborative task are likely to possess different kinds of knowledge and so will engage in interactions that will allow them to pool these varying resources to accomplish their tasks. In addition much knowledge is shared by the individuals, which enables the adoption of various communication practices (Rogers, 1997).

Distributed cognition thus put forward that cognitive artefacts amplify the cognition of the artefact user. This basic concept of distributed cognition has been proved not only by Hutchins' study of the use of monograms in ship navigation, but also by Rogers in the context of engineering practice (1992; 1993), Hutchins and Klausen's (1996) study of cognition in the cockpit, and Halverson's (1995) study of air traffic control.

The distributed cognition approach uses a number of methods: from detailed analysis of video and audio recordings of real life events, to neural network simulations and

laboratory experiments. The type of methodology adopted depends on the unit of analysis that is being adopted and the level at which the cognitive system is being explained. Interaction of various actors and artefacts is unique to distributed cognition. Misnky & Papert (1987) point out that symbols of interaction can be spotted in the boundaries of the various units of analysis, as the flow of information is relatively less there.

2.6 Conclusions

This thesis is written with the aim of laying the foundation of user-centred roads. To do this, an essential relationship between SER and SA needs to be established. Therefore it is imperative to choose a suitable model to explain SA in a road transportation context and link it with SER later in this thesis. After reviewing the literature it can be concluded that the concept of SA is still a challenging one. There are no universally accepted definitions of SA nor there is a SA model accepted by researchers and practitioners alike. This finding is also supported by a recent review of SA by Salmon & Stanton (2013). In the context of driving much attention has been paid to improve SA in driver training and education. This means that literature is concentrated on improving the SA of car drivers, rather than that of roads. The latter is also seen through this review.

Two perspectives (individual and computational) do not provide the bigger picture of a situation, but rather two extremes of a situation. Thus, the interactionist perspective's DSA model acts as a bridge between the two, and is the best and the latest model to explain SA in socio-technical systems such as roads. Distributed situation awareness as given by Salmon et al. (2009) is a model that presents the interactions of agents (humans and non-humans) within a subsystem. This approach helps to promote a better technology-mediated interaction in systems and its agents. This perspective or model proposes that within a system each agent holds its own SA which may be different from that of other agents (Stanton et al., 2006). SA of each agent is based on the goal or purpose thus sharing of this SA is rarely seen (Stanton et al, 2006). Unlike the previous two perspectives (psychological and computational) this perspective views SA as a collaborative or dynamic process instead of a product, binding agents together on a moment-to-moment basis. DSA is a system-oriented rather than individual-oriented approach, thus significantly differing from the models of Endsley or Salas et al. This

implies that knowledge underlying DSA is distributed across the system (Stanton et al., 2006). Furthermore, this model supports implicit communication of information through feedback, rather than a detailed exchange of mental models.

The popular Endsley's model and others which stem out of it are based on poorly explained or poorly researched ideas like mental models. In this regard, DSA is very robust. Unlike the other two perspectives based on one theory to support them, DSA is based on schema theory, Neisser's perceptual cycle (1976) and distributed cognition (1995). According to schema theory, the individual hold schemata which are activated by the task which is being performed (Stanton et al., 2011). Through task performance the phenotype schemata are created through the interaction between people, world and artefacts (Salmon et al., 2009). According to the perceptual cycle, SA resides in the cycle of activity rather than in one operator alone, and is associated with agents but does not reside with them as it is borne out of the interactions between them (Smith & Hancock, 1995; Salmon et al., 2009). Thus, it is fair to conclude that DSA is an emergent property of a socio-technical system which is characterised by distributed cognition (Hutchins, 1995) and arises from the interaction between people and environment. This essentially forms the crux of the thesis and highlights its appropriateness within the human factor community.

It is clear from Chapters 1 and 2 that roads do not work in isolation from the rest of the elements on them. Individuals are so closely coupled with their environment that they cannot be analysed in isolation (Bedny & Meister, 1999). On the contrary they work in close contact, where people and environment form a joint cognitive system (Hollangel, 2001) where every agent has its own unique role to play. This is echoed by Gorman et al. (2006) who consider SA as an interaction-based phenomenon. The roles that agents play are in direct proportion to the goals they seek to achieve. Road accidents occur if there is an incompatibility between these goals (Salmon et al., 2011). They maintain this compatibility by interaction between human and non-human agents, and this interaction is essentially in the form of feedback. This idea is congruent with the three theoretical underpinnings of DSA (perceptual cycle, distributed cognition and schemas). Moreover, it is only DSA which has the ability to highlight the interactions between engineering elements and human road users. Hence, DSA will be used in the rest of the thesis to understand the role of various road elements in everyday driving. Concept mapping and

propositional networks will be used to explore this relationship and information transfer between infrastructural and social elements.

To understand the relationship between the various agents of a system it is essential to take a systems perspective and DSA forms the thrust of that perspective. Stanton et al. (2006, p.1291) define distributed SA as "activated knowledge for a specific task at a specific time within a system". Though the model has been conceived very recently, its applications are very diverse and generic and appear to be the ideal model for explaining the road environment and other members in it.

The central hypothesis of this work is concerned with exploring the relationship between SER and SA. As mentioned in the introduction there are a number of shared principles between the two. Furthermore, a well-designed SER plays a critical role in supporting SA for different road users and helps them stay connected (Salmon et al., 2013) thus promoting cognitive compatibility. SER can also go further than only activating schemas but can create new schemas or knowledge structures. For instance, a non-SER intersection cannot trigger schemas for car drivers to have a lookout for motorcyclists and cyclists (Salmon et al., 2013). Car drivers at intersections of non-SER tend to only focus only on traffic lights, speed and road environment ahead (Salmon et al., 2013). A SER intersection on the other hand will have a dedicated cyclist lane. This will create two new schemas for car drivers. They are; anticipating pedestrians who might be walking on the cyclists' lane thus more heterogeneous traffic. Second schema that this will create is the decreasing speed to accommodate cyclist negotiate the intersection.

Therefore this thesis is written to make a meaningful extension to the already existing SER research. Chapter 2 is concerned with the theoretical aspects of DSA, but the chapters to come will look into more experimental aspects. To make DSA a meaningful part of SER a systematic approach is adopted. First is choosing a mixed method approach (Chapter 3) then detecting the source of driver SA (Chapter 4). The methodology chosen will be put to test in a pilot study. Following the satisfactory results of a pilot study, a large scale study to assess DSA will be undertaken (Chapter 5). To make a strong association between SER and SA, elements of SER will be derived

from DSA methods (Chapters 6 and 7). Finally, the implications of the SER-SA relationship will be discussed in Chapter 9.

Chapter 3 : Review of Measurement of Situation Awareness Measurement Techniques

3.1 Introduction

An important conclusion from Chapters 1 and 2 is that there is a tight coupling between human elements and the road environment. Roads, if designed to be cognitively compatible can reduce the accident toll. Roads contain objects and design elements which can communicate with drivers to help them drive safely. A well-designed road will have these explicit communication links which will enhance 'road readability' (Mazet et al., 1987). Given this complex nature but the pivotal role of road in driving it is essential to determine a methodology which can reveal this communication between driver and road. Furthermore, it is also essential to determine which objects on the road contribute more than others in drivers' perception.

From the first part of the review it can be concluded that SA is now a well-established construct and its importance has been well-recognized in system design and assessment (Endsley et al., 2003; Salmon, 2008; Shu & Furuta, 2005). This puts enormous pressure on researchers, practitioners, and procedure designers within the system to evaluate individual and team SA in these environments. Accurate measurement of SA is not only critical to theoretical advancement of SA but also to system design, design for training programs, organizational structure and channel evaluation efforts. Researchers need valid and reliable methods of assessing SA in order to test and advance SA theory, whilst designers and practitioners need means of assuring that SA is improved and not degraded by new artefacts, systems, interfaces, procedures or training programmes.

The purpose of this chapter is to compare and contrast the different SA measurement approaches that are available to researchers and practitioners undertaking an assessment of the construct. The chapter will start with providing a concise introduction to measurement in ergonomics, followed by problems in determining reliability, validity and sensitivity of SA methods. Finally, the chapter will conclude by defining a

methodological approach to be followed for the rest of this thesis. Together with Chapter 2, Chapter 3 also fulfils the first objective of this PhD.

3.2 Introduction to Measurement in Ergonomics

"Ergonomics (or human factors) is the scientific discipline concerned with the understanding and interactions among human and other elements of the system, and the profession that applies theoretical principles, data and methods to design in order to optimise human well-being and overall system performance."

- International Ergonomics Association (IEA), as cited in Wilson & Corlett (2005, p.4-5).

This definition from the IEA highlights the functional nature of ergonomics, its implication in design and the interaction of humans with other elements in the system. The implication of ergonomics in design has led to the concept of user-centred design (UCD). UCD is a design process in which the end users decide how a product should be designed (Abras et al., 2004), and the term was originally coined by Donald Norman in 1986. Central to the UCD is usability; "a product or service is truly usable when a user can do what he or she wants to do, the way he or she expects to be able to do it, without hindrance, hesitation and questions" (Rubin & Chisnell, 2008, p.4). Thus, in order to provide a user-centred design, the first step is to determine the needs of the user so as to integrate those needs into the system or product and produce a UCD. Poorly designed products are very quickly negated by people and this has laid more emphasis on devising more innovative methods to gauge the requirements of users. To do this, adequate human factor methods are required.

Stanton and Young (1997; 2003) pointed out that it is the usability aspect of ergonomics which has triggered research in ergonomics methodology. Since UCD is taken seriously by consumer-driven economies, a plethora of human factor methods are now available to the human factor researcher and practitioner. Major works concerning human factor methods are Wilson & Corlett (2005); Stanton & Young (1999); Diaper (1989) (as cited in Stanton & Young, 1997); Corlett & Clarke (1995) (as cited in Stanton & Young, 1997) to name a few. Stanton et al. (2013) also produced an excellent compendium of over 200 ergonomic tools specifically for cognitive ergonomics. This

rapid growth in research on human factor methods focussed on quantity rather than on quality, or as Stanton & Young (1997) put it, "pragmatic development over scientific rigour". Yu et al. (2002) contradict this finding of Stanton & Young (1997) of the pragmatic growth of human factor methods.

According to Yu et al. (2002) the development of measurement tools has not kept up-to-date with the theoretical development of the subject. Yu et al. (2002) further cautioned that lack of development of human factor methods may result in stunted growth of the subject. Although, Yu et al. (2002) and Stanton & Young (1997) differ in terms of growth of methods, they do agree on lack of objectivity of the traditional human factor methods.

This range of different human factor methods provides a problem of choice to human factor practitioners and researchers (Stanton & Young, 2003; Stanton & Young, 1997). For example, Wilson and Corlett (2005) grouped methods into 5 categories. These are: general methods, methods of collection of information about people, system analysis and design methods, human-machine system evaluation methods, methods to assess demands and effects on people, and methods to develop ergonomic management programmes. According to Wilson and Corlett (2005) each of the methods in every group namely suffers from three debates innate to human factor methods: quantitative vs. qualitative, objective vs. subjective and field vs. laboratory study. Furthermore, researchers and practitioners also have to evaluate some practical issues related with each method, for instance training time, administration time and benefits of the chosen methods (Stanton & Young, 1997). Thus choosing from a wide range of available methods and distilling them through these debates makes selecting a human factor method a cumbersome process.

No practitioner or researcher will be willing to use any method until its effectiveness is proven or documented. Yu et al. (2002) suggested that traditional human factors methods (like task completion time) suffer from lack of objectivity, sensitivity, and theoretical weakness. For example, eye movement and verbal protocols offer good scientific insights into the user, but are extremely time-consuming and subjective to analyse (Yu et al., 2002). Thus, Yu et al. (2002) concluded that traditional methods fall short of satisfying the four criteria of objectivity, quantitative approach, theoretical

strength and sensitivity. Young & Stanton (1997; 2003) on the other hand put forward that weakness of psychometric properties of the existing human factor methods may be a restrictive factor to use. This is also consistent with Annett's observation, of no mention of reliability and validity of ergonomics methods in their manuals (Annett, 2003). As Wilson and Corlett (2005, p. 22) emphasise, "the proof of the pudding is in the eating", meaning that the best or the most objective way to ensure if the method is really useful is to determine its reliability and validity which are the two cornerstones of psychometrics. Therefore, a human factor method should be able to solve a human factor problem which lies at the interface of human behaviour and engineering. A human factor method can be deemed scientifically robust only if it is sensitive, as well as rich in reliability and validity.

Psychometrics, a frequently occurring term in personality measurement and assessing cognitive abilities of an individual, are seldom mentioned in the domain of ergonomics (Annett, 2002). Surprisingly, Stanton & Young (1997; 1998; 1999) are the first to discuss this explicitly (Annett, 2002). In a review of major ergonomics methods literature (e.g., Wilson & Corlett, 1995; Diaper, 1989; Jordan et al., 1996; Karwowski, 2001; and many more) by Young & Stanton (2003), they found that there was no mention of validity for any of the ergonomic methods addressed. Stanton & Young (1999) referred to this as the "failure of ergonomists to live up to their proclaimed standards of scientific probity" (cited in Annett, 2002, p. 2).

3.2.1 Reliability

Joppe (2000, p.1) in Golafshani, 2003 p.598 defined reliability as, "…the extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable." Annett (2003) in very simplistic terms defined reliability as the repeatability of results over time by same or different examiners in the same or different conditions. In numerous psychological texts, reliability is referred to as the consistency of results. There are three types of reliability as given by Miller & Kirk (1986):

1) *Parallel form reliability* assesses the consistency of results of the same construct by using same or different forms of the same test to same or different group of respondents.

2) *Inter-rater reliability* assesses the consistency of results with respect to the degree of concordance or agreement amongst different raters measuring the same construct at same or different times.

3) *Internal consistency reliability* confirms the homogeneity of the items in a test. This assesses the consistency of the results by assessing the extent to which the items in a test are measuring the same construct.

Charles (1995) later added one more type of reliability called test-retest reliability. If an individual has same score on a test at two different times then a test is said to have high test-retest reliability.

Human factors methods rely on inter-rater reliability and to some extent on test-retest reliability. This is due to the fact that human factor rating scales do not have different forms of a same test. Additionally, internal consistencies of none of the tests are recorded. The first record of psychometrics of human factor methods by Stanton & Young (1999) refers to reliability and validity of almost 60 human factor methods with no explicit reference to specific types of reliability. This indicates the uncontrolled quantity-focussed growth of human factor methods. Additionally, the nature of human factors as a discipline is such that its reliance on classroom tests are highly criticised (e.g. disadvantages of subjective rating scales are discussed by Annett (2002)). Stanton and Young (2003) demonstrated variation in results of human factor methods when applied by experts and when applied by less trained researchers. Patrick et al. (2000) reported training and experience to be directly proportional to increases in quality of task performance. As human factors methods depend on training of the researcher, it is extremely effortful to determine an objective Figure for inter-rater reliability. Moreover, data collection in order to design a method is always subjected to bias due to inadequate or biased sample size (Annett, 2002). This results in unsatisfactory reliability of human factor methods.

However, determining the reliability of human factor methods seems to be virtually impossible, but it is not in reality. An attempt should be made to develop more

structured methods as these methods can minimise biases and deny individual limitations (Annett, 2002). Diaper (2001) referred to the use of analytical methods (like Hierarchical Task Analysis, Task Analysis, error analysis methods, and methods used to understand the working of any complex system) as more of a "craft skill" than a structured method.

Baber & Mirza (1996, as cited in Stanton & Young, 1998) specified that the frequency of use of a human factor method depends on its ease of use and accessibility. They found that observations, heuristics (speculative formulation based on experience with intuition serving as a guide), questionnaires, interviews, and checklists are used repeatedly in combination or in isolation to evaluate a product. All of these are based on are subjective assessment, and not well-structured techniques. This indicates the need to make objective human factor methods accessible and easy to use with less training time

3.2.2 Validity

Validity is defined as the extent to which a test measures what it purports to measure (Miller, 1986). Annett (2002) simplistically deems a method to be valid if it sustains all objections, scrutiny and contradictions. Validity has arguably been regarded as the most important concept of psychometrics (Lissitz, 2009). Lissitz states three important characteristics of validity: it is not embedded in a test, it refers to the interpretations of actions made on the basis of scores on the test, and it must be determined with reference to what the test is measuring and how it is doing so. Lee Cronbach (1954) listed four different types of validity:

1) *Construct validity* refers to the extent the test measures the trait or theoretical construct that it is intended to measure (Miller, 2001). Lissitz (2009) claims construct validity to be the most important form of validity whereas Rosenthal & Weston (2003) regard construct validity as the one of the most important subjects of psychology.

2) *Concurrent validity* describes the correlation of a new test with the existing test that claims to measure the same construct (Rust & Golombok, 1999).

3) *Predictive validity* is a degree of measurement of how well a test predicts future performance on the criterion measured.

4) *Content validity* is the "representativeness and relevance of the assessment instrument of the construct being measured" (Marnat, 2003, p.51).

Face validity was given by Mosier, 1947 is the simplest form of validity. A test is said to have high face validity if its purpose is clear to naïve respondents (Nevo, 1985). Similarly, if a test's purpose is unclear to its respondents it is said to have low face validity.

It is worthy to note that a human factor method which is reliable is not necessarily valid, but a method which is valid is certainly reliable (Young & Stanton, 2003). Thus, the relationship between reliability and validity is unidirectional (Stanton & Young, 2003). Annett (2002) differentiated between construct and predictive validity in terms of human factor measurement tools. To do so, Annett (2002) distinguished human factor into analytical and evaluative methods. Analytical methods (e.g. task analysis, error analysis) are used to understand the working of a complex system and evaluative methods (e.g. workload, fatigue) are used to measure a parameter of the system. Analytical methods are to display high construct validity and evaluative methods are supposed to show high predictive validity (Annett, 2002). Furthermore, Annett (2002) suggested that validity of analytical methods depend on its underlying theory. This implies that an error analysis method will look very different if it is based on Freudian theory (Annett, 2002). Stanton et al. (2013) argued against dichotomy of methods, and proposed that human factor methods are not mutually exclusive. It is possible for a human factor method to be reliable, valid and sensitive at the same time (Stanton et al., 2013).

Kanis (2010) in his review of papers from *Ergonomics*, *Applied Ergonomics*, and *Human factor* concluded that the terms validity, validation and valid have been used interchangeably by authors (Kanis, 2010). This has blurred the scientific usage of the term (Kanis, 2010) in the human factor community. Moreover, Kanis (2010) adds that it is impossible to achieve the concept of validation as it is based on counter-evidence. Validation only holds true as long as counter-evidence fails (Kanis, 2010). Lack of validity studies in human factors literature can also be attributed to the abundant human factor methods available to a practitioner or researcher (Stanton & Young, 1998). A method which seems to be useful to one practitioner might not be useful to another

(Wilson & Corlett, 2005). This has restricted the use of most of the human factors methods by its discoverers only (Stanton & Young, 1998).

3.2.3 Summary

To improve the reliability and validity studies of human factor, researchers should express the results of human factors method application in terms of cost-effectiveness (Stanton & Young, 2003). This means quantifying the cost-benefit ratio (Young & Stanton, 2003). The cost of applying the method can be calculated in terms of required persons, application time, materials, etc. and subtracting this from the estimated savings generated by improved design. The net Figure is proposed as the benefit brought by using the ergonomic methods. Additionally, cost-effectiveness is said to be linked to the reliability and validity of a method. In a survey of human factors methods Stanton and Young (1998) reported that most practitioners were unaware of any reliability and validity data regarding the method in hand. Proving the utility i.e. the cost-effective nature of ergonomics methods can improve this condition and can help to multiply studies in reliability and validity of methods in ergonomics. It is noted by Stanton and Young (2003) that the absolute value of validity cannot be calculated in ergonomics. This is due to the reason that the decision of the user is influenced by the way in which information is presented to the user (Kanis, 2010; Stanton & Young, 2003). Thus, complete information about the getting the data and how it is interpreted (Stanton & Young, 2003) are required in order to make substantial contribution to validity studies of ergonomics.

Despite the above problems in achieving a substantial standard of ergonomics methods, many researchers have put forward few criteria against which a human factors method should be evaluated before being put into use. Wilson and Corlett (2005) gave eight criteria: validity, reliability, generalizability, non-reactivity, sensitivity, feasibility of use, acceptability, and ethics and resources. It is rarely possible to satisfy all eight criteria (Wilson & Corlett, 2005). Stanton et al. (2013) have evaluated over 200 human factor methods against training time, application time, tools needed, reliability, validity, related methods, type, advantages, and disadvantages. The criteria of Stanton et al. (2013) are too detailed to evaluate basic and physical ergonomics methods. For the purposes of this PhD, methods will be evaluated against reliability, validity and

sensitivity. Furthermore, a multi-method approach will be adopted in order to achieve the objectives of the study. A multi-method approach will give the liberty to mix from qualitative and quantitative methods and in-field and laboratory techniques (Wilson & Corlett, 2005). This is also in line with Wilson & Corlett's (2005) recommendation of a triangulation approach for overcoming the disadvantages of one method with the advantages of another. For example, simply questioning and asking the operator to rate their responses in complex systems may yield incomplete information. However, combining rating with concurrent verbal protocol can provide insights into the decision-making of the operator in the system.

3.3 Why is measuring SA necessary?

Before exploring the causes which make SA measurement difficult, the intermediate section will shed light on why SA measurement is necessary.

Measuring situation awareness is essential due to its widespread applications in design across all domains. Despite being first recognized in aviation it is used widely in everyday tasks like car driving. According to Sonnewald et al. (2003) individuals must do more than just perceive elements in the environment. Endsley (1995) suggested individuals should integrate their perception of the situation in light of the goals they have. This is consistent with the definition of SA put forth by Dominguez et al. (1994, p.11), as the "continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture in directing further perception and anticipating future events". Thus, it can be concluded that SA presents an integration of individual as well as social levels or cognitive orientation (Dominguez et al., 1994). This very property makes measurement of SA all the more necessary and challenging.

Salas & Bowers (1997) identified three important reasons to assess SA. The first reason is inextricably related to the development of theory beyond a conceptual level (Salas et al., 1995). Salas and Bowers (1997) affirm that without the existence of robust psychometric measurement tools, it is impossible for a model of SA to progress beyond an intangible level without substantial development in its measurement. This is surprisingly similar to Annett's (2002) reasoning for lack of validity studies in

ergonomics. Annett (2002) proposed that validity of analytical methods such as task analysis is in direct proportion with the development of its corresponding theory. Hence, it can be concluded that advancement in theory and measurement in the context of SA go hand in hand. In the case of SA, debate in one will lead to debate in the other.

Secondly, without quantifying SA requirements (Salas et al., 1997) it will be almost impossible to design to support SA. Due to the demand in user-centeredness of a product, there is increasing pressure on the designer to know what the user expects from a product and what his limitations are. All of this information also plays a pivotal role for providing feedback of performance during training (Salas et al., 1997). Finally, Salas et al. (1997; 1995) suggest measurement of SA will help to form effective training programs. Measures scoring high on psychometrics will serve as an indicator of the effectiveness of training program (Salas et al., 1997). On the basis of these necessities to measure SA, Salas et al. (1995) provide a "what", "when" and "how" of assessing SA.

Although SA as a construct and its measurement is very important, this measurement is not simple. Assessing SA causes too many problems for experts on the subject. Some of these problems have been explored in the consequent section. One of the conclusions of Chapter 2 (Situation Awareness: Literature Review) is that there is enormous contention around the theoretical aspects of SA. As pointed out earlier there has been an exponential growth of research in in SA since its inception. Stanton et al. (2009) classified the three distinct ideas of SA into three different perspectives. A plethora of models and definitions is available to SA practitioners. For instance, Bell and Lyon (2000) refer to SA as a product of cognitive processing whereas Smith & Hancock (1995) define SA to be externally-directed consciousness. Researchers further state that every model has an element of truth in them (Stanton et al., 2006). To summarize, this incongruent literature and lack of an objective definition poses a threat to the attainment of *construct validity* as there is no agreement on what the construct actually is. Moreover, this also makes the measurement of SA far more difficult than it should be. To make matters worse there also exists a debate between the nature of SA being a product or a process.

Almost all the measures of SA are developed in line with a model of SA (Salmon et al., 2009). As noted from Chapter 2 of this thesis, all models of SA have different

theoretical underpinnings. For instance, Endsley's model has the backing of mental models and information processing, whereas DSA is theoretically pinned by Neisser's perceptual cycle, schemas and Hutchins' distributed cognition. Thus unless the human factor community has reached a universally accepted theory of SA there will always be contention surrounding the measurement tool of SA. This is also supported by Annett (2002). According to Annett (2002), HTA will look very different if it is based on Freud's theories. Measures of SA are developed based on the theoretical underpinning of the aforementioned models in Chapter 1. A method based on a model like that of Endsley's will show less *sensitivity*. This is due to the fact that Endsley's model is itself based on ability, experience, training, long-term memory stores, etc., and that these elements of cognitive psychology lack a sensitive measurement tool themselves. SAGAT (Situation Awareness Global Assessment Technique) is closely linked to Endsley's model of SA, and presents a classic example of problem of sensitivity. Under SAGAT subjects are unexpectedly stopped during their performance in a simulator. The subjects are then asked questions, testing awareness of the current situation (Salvendy, 2012). Recalling disparate elements will deem an individual to have less SA. SAGAT basically probes the individual to yield specific information relevant for Endsley's SA. Moreover, it indirectly assesses attention rather than just SA. SAGAT measures memory, attention and SA instead of focussing on SA.

Human factor methods suffer from the problem of validity (Stanton & Young, 1997). Unfortunately, SA methods are not untouched by this problem of *validity*. In the case of SA, the terms validity and sensitivity are often used interchangeably. For example, Salmon (2008) referred to a valid SA measure as, 'the one measuring SA and only SA and no other psychological product or process' (p.43). This is basically how sensitivity is defined. Validity is the measure of the degree a method measures what it purports to measure. Validity as noted by Stanton & Young (2003) is mostly assumed rather being tested. Tests are often assumed to be valid if they are found to be reliable; however, as pointed out before, the relation between reliability and validity is often unidirectional. In an exhaustive review of over 90 SA methods conducted by Stanton et al. (2013), validations of not all methods were found. This is probably due to the reason that SA methods, like human factor methods in general, are often used by their inventors only and not by other experts of the subject.

It can be inferred that SA measurement is a problematic task due to the very basic abstract nature of SA itself. It is also difficult to measure as the term is mixed with other related constructs. Lack of a universal definition and model also acts as a significant hindrance towards objective measurement of SA. SA measurement suffers from all the problems which are experienced by human factor methods in general, with the addition of a lack of acceptability of one single definition with theoretical underpinnings for SA.

The number of validity studies in the context of SA measurement is still not enough, so to make validation studies a basis for selecting SA methods would not be wise. One of the secondary objectives of this PhD is to inform the current theory in SA and contribute to the development of SA for the evaluation of SA in complex interactive systems, particularly related to those of interaction in safety-related environments such as road driving. One of the criteria for selecting a method will be its ability to measure SA of stationary, non-living objects. Another criterion is *ecological validity* (Brunswick, 1952) so that it can be used in safety-critical domains. Ecological validity is the metric of correspondence of the experimental environment with the real environment. These criteria are to be combined with Endsley's properties of what a SA measurement tool should have. These are the ability to measure exclusively, and high sensitivity such that the technique can accurately detect changes in SA caused by alterations in technologies, while measures for SA remain unaltered during measurement.

3.4. Review of Situation Awareness Methods

The measurement of SA in the context of roads poses a substantial challenge to the human factor community. Roads work as a rich, complex, socio-technical team of humans, environments, and vehicles, etc. The team members of a road are dispersed across various geographical locations. Inaccurate actions may result in disastrous and sometimes fatal results. Measurement of SA in a road context should be such that it can assess SA of different agents scattered across different geographical locations simultaneously (Salmon, 2008). Alternatively, it can attempt to calculate SA of each road type and apply this to different users across a country. This thesis aims to seek the latter.

Table 3.1 Summary of Different SA Classification Schemes

Authors	Year	Categories of Classification
Sarter & Woods	1991	Subjective measures, query techniques, implicit performance
Fracker	1991	Explicit, implicit and subjective ratings
Gravel & Schopper	1994	Retrospective, concurrent, subjective, process, performance, signal detection theory
Adams et al.	1995	Online probe, indirect probes, and model based
Durso & Gronlund	1999	Online vs. offline measurement, direct vs indirect measurement
Endsley	2000	Process, direct, performance and behaviour measures
Sandom	2001	Physiological, performance, subjective, questionnaire
Stanton et al.	2013	SA requirement analysis, freeze-probe, real-time, self-rating, observer-rating, distributed SA techniques
Salmon	2008	SA requirement analysis, freeze-probe, real-time, self-rating, observer-rating, performance measures, process measures, distributed SA techniques

The aims of reviewing SA methods are three-fold. First is to choose an appropriate method or a combination of methods to fit smoothly with the aims of the current PhD. Second is to shed light on the range of SA methods available to a practitioner or researcher. Finally, investigate the degree to which individual perspectives on SA have influenced SA measurement tools. A concise literature review suggests existence of different classification schemes. A summary is presented above in table 3.1.

One of the reasons so many classification schemes exist for one single construct is the disagreement among the practitioners and researchers whilst classifying methods. Such a disagreement is caused due to a lack of universal definition of SA which in turn hinders high construct validity for methods. It is like a vicious circle. Human factor methods themselves do not have any widely agreed upon criteria for classifying methods – e.g., Sanders & McCormack (1993) and Meister (1985) follow different criteria for categorizing ergonomics methods.

Following McCormick & Sander, 1993 this research investigation will satisfy two requirements: practical and psychometric. Practical requirements have been given considerable importance in human factor literature prior to 2000 with the exception of Stanton et al. (2013). Common practical requirements of Meister (1985) and McCormack and Sanders (1993) are objectivity, ease of use, and cost. Psychometric requirements are: reliability, validity, freedom from contamination, and sensitivity (McCormack & Sanders, 1993).

In the case of SA, Uhlarik & Comerford (2002) stressed on focussing on face, concurrent, predictive and construct validity while considering SA methods. This is consistent with the Salmon et al. (2008, 2006) assessment of SA methods. In the case of this PhD, three forms of validity i.e. face, predictive and construct validity will be considered. Concurrent validity refers to how well a test correlates with its previous versions which have been validated (Miller & McIntyre, 2005). Concurrent validity will not be taken into account in this thesis because of a lack of SA measures available, and a lack of SA measures designed specifically for measuring team SA. Thus there are not many team SA measures to correlate with.

Endsley (1995) put forward three properties which a SA method must have. These are:

1) The ability to measure SA and only SA. This is in other words is construct validity

2) Sensitivity to alteration in new technologies and design.

3) The ability to measure SA unaltered.

Taking into account the various contentions in terms of measurement of SA, SA methods in the subsequent section will be evaluated for cost, ease of use and administration time. It will also be checked for inter-rater reliability, validity and sensitivity

SA methods have been widely grouped under nine broad categories for the purpose of this thesis. These are:

1. SA Requirement Analysis
2. Freeze-probe Techniques
3. Real-time Probe Techniques
4. Self-rating Techniques
5. Observer Rating Techniques
6. Performance Techniques
7. Process Indices
8. Team SA Measures, and
9. Naturalistic Studies.

A brief explanation of each category is as follows.

3.4.1 SA Requirement Analysis (Endsley, 1993)

Type: Task Analysis

Description: According to Endsley (2001, p.8) SA requirement analysis can be defined as "those dynamic information needs associated with the major goals or sub goals of the operator in performing his or her job". Endsley (2001) further added that SA requirement analysis does not only provide an insight of perception of elements in the environment but also throws light on how that perception is used to form a decision. Along similar lines Salmon et al. (2008, 2006) reported that SA requirement analysis goes beyond merely defining information needs of a road user but it also explains how that information or knowledge interact with each other for the driver so that he can

make sense of the environment. Matthew et al. (2004) state the importance of SA requirement analysis by directly linking it with system design. Therefore, determining SA needs of the agents within a system forms the first step of design. Knowledge of SA of the system helps an engineer to design a system which maximises human performance, rather than sufferings from information overload (Matthew et al., 2004). SA requirement analysis is generic in nature but has a very high administration time. There is no data available regarding the reliability and validity of this method (Stanton et al., 2013). However, due to the subjective nature of the tool, establishing inter-rater reliability may be useful for scientific robustness.

Advantages:

1. It is generic in nature and has been already been applied in air traffic control, aviation and military domains (Stanton et al., 2013).
2. SA requirement analysis specifies the knowledge required for SA of task under analysis (Stanton et al., 2013).
3. SA requirement analysis if conducted by experts can ensure high levels of validity.
4. The result of SA requirement analysis informs designers aiming to optimise SA of the system through design.
5. Results are also useful for developing SA measures for a particular task (Stanton et al., 2013).

Limitations:

1. Reliability and validity of the method are difficult to establish (Salmon et al., 2010).
2. It is a very lengthy process. It is generally meant to be used for complex tasks hence a large number of sub-tasks need to be analysed for each task under analysis. Therefore it is very time-intensive.
3. SA requirement analysis if conducted from a scratch needs to begin with goal-directed task analysis (such as hierarchical task analysis), interviews, and the development and administration of questionnaires. Hence, access to experts is required at all times.
4. The result of SA requirement analysis is dependent on the interview skills of the interviewer (Stanton et al., 2013).

3.4.2 Freeze-probe Techniques

Examples: Situation Awareness Control Room Inventory (SACRI) (Hogg et al., 1995); Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1995); SALSA (Hauss & Eyferth, 2003)

Description: Freeze-probe techniques, as the name suggests, involves freezing of simulations at random during task performance. This means that the screen will be blacked out at any point during performance of the task. During the blackout of the screen, a rating scale or an online probe is administered to test the subject's knowledge of the current situation. Their responses are correlated with the state of the simulator to determine the SA of the operator.

Freeze-probe techniques in general are used more than the other available methods. This contributes to a large number of validation and reliability studies carried out for freeze-probe techniques alone. These techniques offer a direct and objective measurement to operator SA (Salmon, 2008). Although appealing, freeze-probe techniques suffer from the problem of being intrusive to task performance and having low sensitivity. For example, in case of SAGAT, it is questionable if the method is really measuring SA or memory (Salmon, 2008). Freeze-probe techniques such as SAGAT and SART (the Situation Awareness Rating Technique) have to be used in conjunction. SAGAT alone cannot offer an insight into an operator's SA. Furthermore, SAGAT and SART can be directly linked to Endsley's three-level model (Salmon, 2008) so the flaws inherent to Endsley's model for measuring teams are also present in SAGAT and SART.

Advantages:

1. All methods under the category of freeze-probe techniques provide direct measurement of participant SA (Stanton et al., 2013; Salmon et al., 2010; Salmon et al., 2009).
2. Freeze-probe techniques remove the problem of collecting data post-trial (Stanton et al., 2013; Salmon et al., 2010; Salmon et al., 2009).
3. Methods under freeze-probe techniques are used the most by researchers and practitioners.

4. SAGAT is the most successful freeze-probe technique used and therefore has a wealth of validation studies (Stanton et al., 2013).

Limitations:

1. On the downside, freeze-probe techniques are high-cost. All freeze-probe studies require task and system simulation, which is an expensive process.
2. It is intrusive to primary task performance (Stanton et al., 2013; Salmon et al., 2010; Salmon et al., 2009, Salmon et al., 2006).
3. The queries may direct the attention of participants to SA elements in the environment (Stanton et al., 2013; Salmon et al., 2010; Salmon et al., 2009). Much of the methods under freeze probe techniques such as SACRI and SALSA are influenced by SAGAT. This leads to several problems. The first is originality; for example, SALSA is reported to be very similar to SAGAT (Stanton et al., 2013). Secondly, it is only SAGAT which has high validity and reliability, while the other methods lack these as many are still in their infancy (such as SALSA and SACRI). Finally, all of these methods are inspired from SAGAT, therefore they are directly linked to Endsley's three-level model, require simulation, and are not fit for real world use.

3.4.3 Real-time Probe Techniques

Examples: Situation Awareness Present Assessment Method (SPAM) (Durso et al., 1998); SASHA_L and SASHA_Q both developed by Jeannot et al. (2003), with the former using online probes and the latter using a post-task probe method.

Description: Real-time techniques are very similar to freeze-probe techniques. They administer SA-related queries online, but without any simulation of the task under analysis. Classically, self-rating techniques are administered after or during task completion. Since it is a real-time technique, the queries are generated online by experts, without freezing at specific points. Participant SA is assessed from the answer content and response time.

Real-time probe technique has several advantages over freeze-probe techniques. Since it is (typically) administered online with the active task, it is non-intrusive to primary task

performance. Because the nature of administration of probes, real-time probe techniques provide a direct measure of operator SA. Finally, probing participants for their SA allows it to the directly compared with the objective state of the world. Endsley and Jones (2000) claim real-time probes to be a worthwhile option for measuring SA when there is no simulation available.

Advantages:

1. They are quick and easy to use.
2. No simulation is required for the task under analysis.
3. This provides an objective measure of SA.
4. SASHA _L and SPAM are online and telephone methods, respectively. Therefore this removes the problems associated with data collection post-trial (Stanton et al., 2013).

Limitations:

1. SASHA and SPAM have no validation data available (Stanton et al., 2013). They also have questionable construct validity. For example, participant SA in SPAM is assessed on the basis of response time (Stanton et al., 2013). The time taken by the operator is a rough indication of the workload of the participant (Salmon et al., 2009).
2. The measures under this technique are developed by other practitioners and experts of human factor and applied psychology (also referred to as subject matter experts; SME). Obviously generating them while a participant performs a task puts significant pressure on SMEs.
3. Although real-time probe techniques assert to be non-intrusive, Salmon et al. (2008) and Stanton et al. (2013) question the degree to which this intrusion is reduced.

3.4.4 Self-Rating Techniques

Examples: Self-rating techniques is the largest category and contains the highest number of human factor methods. Some of these are the Situation Awareness Rating Technique, or SART (Taylor, 1990); the Crew Awareness Rating Scale, or CARS (McGuiness & Foy, 2000); the Situation Awareness Rating Scale technique, or SARS (Waag & Houck, 1994); the Mission Awareness Rating Scale, or MARS (McGuiness &

Foy, 1994); the Cranfield Situation Awareness Scale, or C-SAS (Demehy, 1997). Of these methods, SART (Taylor, 1990) is the most widely used self-rating technique.

Description: As the name suggests it involves participants providing a score of their own SA via questionnaire, rating scale, etc. Typically, a self-rating technique is administered after task completion.

Advantages:
1. The primary advantages of self-rating techniques are its low cost, ease of administration, and quick processing times.
2. All of the tests under self-rating techniques are fairly generic (Salmon et al., 2009).
3. Since it is administered after the task, self-rating techniques are non-intrusive in nature.

Limitations:
1. Participants are bound to forget periods of low or poor SA when they indicate their response on a rating scale (Endsley, 1995).
2. Self-rating techniques suffer from primacy and recency effects (Stanton et al., 2013). For instance, subjects are usually found to be weak in remembering the beginning of a task. So the scores on rating scales capture the end aspects of a task (Stanton et al., 2013).
3. Since self-rating techniques are administered post-task, sensitivity is questionable. It is likely that the scores represent the working memory of a participant, rather than their SA.

3.4.5 Observer Rating Techniques

Examples: Situation Awareness Behavioural Rating Scale, or SABARS (Matthews et al., 2002)

Description: Like self-rating techniques, observer rating techniques elicit subjective assessment of SA. Experts in the field conduct an observation of overt behaviour of the driver which is likely to constitute their SA. In the case of several agents in a team,

observers/experts provide an assessment of each participant's SA. Summation of each participant SA provides the SA of the team.

The problem of construct validity in SA is very explicitly seen in observer rating techniques. Since no universal definition and model of SA has been adopted by the human factor community, experts have not agreed amongst themselves what SA is. Thus the behaviour they consider to be influenced might not be due to SA, but due to some closely related concept like short-term memory, attention, or hazard perception.

Advantages:
1. They are non-intrusive in nature (Stanton et al., 2013; Salmon et al., 2009).
2. They can be applied during real world tasks and scenarios (Salmon et al., 2009).

Limitations:
1. Despite being a rating technique, it surprisingly has practical limitations which hinder its applications. The first of these limitations is that it relies on one or many experts' observation. There may be cases when experts need to be called back for re-conducting the experiment. This poses many problems and depends on individual availability. This is also a hindrance in establishing reliability. Secondly, team operations like in the case of roads following DSA will require several experts to conduct observations at different places. This also depends on experts' availability.

2. Observer rating techniques suffers from the same limitations as self-rating techniques and observations. These techniques have dependency on performance, whose relationship with SA is not clear. Furthermore, knowing an individual is being observed makes an individual change his behaviour. Finally, experts under observer rating techniques rate their responses in accordance with the behaviour of the operator they see. Stanton et al. (2013) affirm that undoubtedly, external behaviour can provide an idea of operator, however this behaviour is unable to provide precise details of what exactly constitutes an individual's SA.

3.4.6 Performance Measures

Examples: Various

Description: Performance measures can assess SA of particular aspects of a task under analysis (Salmon et al., 2009). Depending on the nature of the task under analysis certain parts of the task are recorded to derive an indirect measure of SA (Salmon et al., 2009). For instance, in the sub-task of negotiating a bend, performance measures can include "collision", "smooth", "merge", "clear". Ma and Kaber (2007) assessed SA of the driver on the basis of number of errors made and consistency of the speed maintained in accordance with the posted speed limit during a simulated driving task.

Advantages:
1. Performance measures are easy to obtain as they are recorded during the task anyway. No extra effort is required to record it.
2. Performance measures don't interfere with primary task performance.
3. These provide an objective measure of SA.

Limitations:
1. Performance measures rely on an unclear relationship between SA and performance. These are based on the fact that high SA will lead to efficient task performance, and vice versa.
2. These are indirect measures of SA.

3.4.7 Process measures:

Examples*:* Eye tracking; concurrent verbal protocol

Description*:* Since the very beginning, physiological measures have always been used to understand experimental psychology more clearly. Human factor methods if used in the right combination yield very accurate results (Wilson & Corlett, 2005). Process measures in this respect provide a very viable option. Process measures essentially involves gauging the cognitive processes of an operator throughout task performance.

A ccommonly used process measure in ergonomics is eye tracking through Facelab. Fixations on certain stimuli inform the experimenter about the location of operator's attention during task performance (Salmon, 2008). Process measures can be very useful for assessing operator SA, and information requirements whilst he is in a multi-agent processes. This is because team process indices can yield how every agent interacts with others.

Advantages:

1. Process indices are cost-effective, less time-consuming, easy to administer and do not require expensive materials.

Limitations:

1. These provide an indirect measurement of SA. According to Salmon (2008) they provide very little information about the product of SA.

3.4.8 Team SA Measures

Examples: Coordinated Assessment of Situation awareness of Teams, or CAST (Gorman et al., 2006); propositional networks (Salmon, 2008)

Description: Chapter 2 noted that there is a dearth of literature describing SA in multi-agent teams. This holds true in the measurement aspect of SA as well. Salmon (2008, 2006) pointed out the scarce attention paid towards the development of team SA measures. SAGAT and SART have been scaled up to assess SA widely.

Originally developed in the military domain, CAST uses changes in task environment to assess team SA. It uses situational roadblocks and the responses towards them to assess team SA. Team SA is measured in the form of coordinated perception and action processes. CAST has been specially developed to assess team SA.

According to Stanton et al. (2006, p.1291), distributed SA can be defined as "activated knowledge of a specific task at a specific time within a system". This activated knowledge is spread across the system amongst human and non-human agents. Propositional networks map this activated knowledge through networking various agents (human or non-human). In a network, the arrows represent the direction of flow

of knowledge or information in a system. This transfer of information is known as an SA transaction. Such a transaction is carried on between agents in a system holding compatible SA, to maintain the SA of the system.

Advantages:

1. The reliability and validity of propositional networks is established through its usage in various domains such as energy (Salmon et al., 2009), military operations (Salmon et al., 2009), rail maintenance (Stanton & Walker, 2011) and road transport (Walker et al., 2013; Walker et al., 2012).
2. Links between knowledge objects are specified in the networks (Stanton et al., 2013).
3. CAST, although a novel approach, provides a picture of interaction between agents and the situation under analysis.

Limitations:

1. The construction and analysis of networks require considerable patience and skills from the analyst.
2. Propositional networks have high application time.
3. The construction of networks is dependent on the data collected through process indices.
4. CAST does not give any indication of individual member SA.
5. CAST relies on disconnect between SA and performance.
6. CAST cannot be applied to real-world tasks.

3.4.9 Naturalistic Studies

Examples: On-road study; on-site study

Description: Naturalistic studies are relatively new for studying everyday driving behaviours. For this purpose, monitoring devices are registered in the operator's car to unobtrusively register driver behaviour such as eye movement, heart rate, etc. (SWOV Factsheet, 2010). In a naturalistic study, the subjects drive the way they normally do without any specific intervention or instructions (SWOV Factsheet, 2010). Such a method provides a very interesting relationship between driver, car and environment (SWOV Factsheet, 2010). This also provides a deeper insight into the driver-

environment relationship which is not elicited by traditional retrospective accident analysis methods.

Advantages:

1. Naturalistic studies are very cost-effective as drivers drive their own cars on a defined course of road.
2. They are not time-consuming and are certainly easy to administer.
3. Naturalistic studies have complete ecological validity.
4. An adequate monitoring device along with a naturalistic study can provide some very interesting insights to the information processing of the operator.
5. It is a recent method therefore its reliability and validity is under construction. However, at the time of writing, naturalistic studies were underway in Japan, Canada and Australia mainly to study driver distraction and inattention.

Limitations:

1. Since the study is done in the natural environment it is difficult to control extraneous variables.
2. It is impossible to repeat the run in the exact weather and traffic conditions.
3. Naturalistic studies need detailed ethical approval.

3.5 Conclusions

The purpose of this chapter was to identify and understand the different measures presented in the literature and to compare and contrast these in order to identify approaches that are most suitable for SA assessing in road transportation context. The following conclusions were drawn from each category of SA methods reviewed.

SA requirement analysis target those "information needs which associated with the major goals and sub goals of the operator in performing his or her task" (Endsley, 2001, p.8). It forms the first stage of system design. It presents the experimenter not only with the information required to complete a task but how it is integrated (Endsley, 2001). In order to define SA of a road it is essential to determine where it is coming from. Therefore, SA requirement analysis will provide a source of SA of the driver in this thesis. Following the result of the requirement analysis a relationship between SER and

SA will be investigated experimentally. The requirement analysis will be done from an already existing and used HTA by Walker (2002) and Walker (2015). This contributes towards the reliability and validity of the analysis carried out. Furthermore, inter-rater reliability will be assessed for 10% of the tasks to increase the robustness of the results.

Freeze-probe techniques despite being very popular and most commonly used are not suitable for this thesis. Freeze-probe techniques are very similar to each other and are mainly derived from SAGAT. These techniques provide an assessment of individual SA, and this is the main reason it cannot be used in this thesis. The preceding two chapters have established the suitability of DSA for road transportation. Consequently, DSA purports that SA of road transport is not the summation of individual SA but the interaction between different agents (living and non-living) on a road. This interaction or SA transaction cannot be elicited by freeze-probe techniques. As pointed out earlier in this chapter, with the exception of SAGAT, psychometrics of freeze-probe techniques are questionable. Despite being rich in psychometrics SAGAT cannot be used because of different variables such as human users, infrastructural elements, and other traffic spread across different geographical location at one single time (Salmon, 2008).

Real-time techniques cannot be used due for three primary reasons. First, these assess individual SA. Secondly, since the thesis is on driving it will require a substantial number of experts to generate real-time queries to the driver. Thirdly, probing the operator whilst he or she is driving may challenge his attention and increase workload making the situation potentially unsafe. In addition the queries generated may divert attention of drivers specifically to the SA elements of the road. Therefore, real-time techniques cannot provide an accurate picture of driver SA on the road.

Self-rating techniques provide SA measures of the participant as perceived by the operator. Drivers drive their vehicle on the basis of the perceived image of the road. Therefore, self-rating technique is used to have an idea what or how drivers perceived the road they drove on. SART scores will be interpreted to understand how aware drivers felt on a particular road. SART is a logical choice from the pool of self-rating techniques as it is backed up by numerous validation studies. Additionally, like other self-rating techniques SART is non-intrusive, quick to administer and easy to use.

Observer rating techniques are not suitable for use in this thesis. The foremost reason for this is because observer rating techniques require repeated access to experts. This is particularly difficult for a dynamic, complex task like driving. Furthermore, observer rating techniques are based on the SA-performance relationship which still remains unclear. Finally, observer rating techniques have an element of subjectivity with different experts interpreting the scenarios under analysis.

Process and performance measures are used to determine road SA. Concurrent verbal protocol analysis, or VPA (Walker, 2004), is used as a process measure to gain insight into driver information processing while driving. Through VPA results, SA can be understood as a process. As the driver drives along the road his verbal protocol varies. This means his information processing is different at various parts of the road. This is of particular importance to this thesis. It is through analysis of VPA knowledge objects in different areas and categories of the road can be identified. Careful analysis can also yield agents holding compatible SA. The above analysis can be done with propositional networks. These networks, apart from providing a list of knowledge objects and concepts underpinning driver SA, also provides how these are linked. Knowing this linkage between different knowledge objects is at the heart of DSA. CAST on the other hand cannot be used, as this method only considers the situation under analysis and the agent.

On-road naturalistic study offers 100% ecological validity and has been found to be very effective in detecting driver inattention (Dingus et al., 2006). Moreover, it offers the experimenter flexibility to use different measures to assess driver SA. Naturalistic study also uses everyday drivers to measure driver behaviour in their natural settings. Thus unaltered driver behaviour can be studied through on-road naturalistic studies. Therefore, an on-road study will be conducted in this thesis.

All methods reviewed to this point show a trend. With the exception of process indices, propositional networks, and naturalistic studies, all other method categories show a heavy influence of Endsley's three-level model. The pros and cons of Endsley's model are discussed in Chapter 2. The review also explains the challenge of determining the psychometrics (reliability, validity and sensitivity) of SA assessment methods. A key conclusion from this chapter is that no one method can serve the purpose of this thesis.

In other words, every method has its flaws and cannot provide an insight to driver SA and how it is formed. An approach is needed which can encompass the advantages of the aforementioned methods by minimizing their limitations. This is consistent with Wilson and Corlett (2005) who suggested that a mix of human factor methods can provide deeper understanding of various ergonomic processes.

Chapters 1, 2 and 3 provided an introduction to thesis, literature review and a concise review of SA methods, respectively. Now the thesis will go into the experimentation phase using the results of Chapter 3. Chapter 4 will start with determining the SA requirements of a car driver. This will be done theoretically through SA requirement analysis, and practically through on-road study. The latter will also act as a pilot study. Following this, Chapter 5 will explain the setup of the naturalistic study. It will detail the process measure(s) to be used, the number of participants, and how data analysis will be done. Chapters 6 and 7 will discuss the results of naturalistic study. Chapter 8 will validate the result of Chapters 7, whose implications will be discussed in Chapter 9.

Chapter 4 : Does Situation Awareness Map Onto Self-Explaining Roads

4.1 Introduction

The thesis until now has focussed on the concept of situation awareness as a whole and identified an SA theory (Distributed Situation Awareness). Furthermore it identified a toolkit of measures such as requirement analysis, naturalistic studies and SART that can help to explain the relationship between drivers and the road environment. This Chapter is about discovering what drivers' SA requirements actually are, and the role of road design within this. To do this, two forms of analysis are undertaken. The first is a requirements analysis whereby a task analysis of driving is used to analytically prototype the information needs of drivers in relation to what information is needed to perform all aspects of the driving task, how much of that information is derived from the road infrastructure, and the modality by which it is presented to drivers. This analytical model is then combined with an on-road pilot study. Here the sources and modality of feedback in a naturalistic environment are explored, and experimental methods to be scaled up in subsequent chapters are tested. The conclusions of the analysis feed into the naturalistic study discussed in Chapters 5, 6 and 7. Furthermore, Chapter 4 fulfils the second objective of this thesis which is to determine ideal source of driver SA and compare how driver's information needs match the results gained from real world data collection.

4.2 Study 1: Situation Awareness Requirement Analysis

The concept of SA has witnessed rapid maturation and has diffused seamlessly in the safety research community. However as noted in Chapter 2 a notable exception in this maturation is the area of road safety and road design (Salmon, 2008). Later research established SA to be critical to road safety due to two reasons. First, Gugerty (1998, p.498) points out that "errors in maintaining SA are the most frequent cause of errors in real time tasks such as driving". Secondly, poor SA can be attributed to more accidents than improper speed or technique (Gugerty, 1997). Inattention which is the key cause of accidents as reported by Treat et al. (1972) is also a by-product of poor SA in the road

and ineffective SER design. Hence, SA is an important factor in driving safety. To design a system with inherent safety it is essential to determine where SA resides and what it is within the driving context. This is known through SA requirement analysis and thus forms the first step towards designing a SA rich system. SA research has offered a set of 16 core design guidelines by Salmon et al. (2009). These principles can also be understood as requirements of a drivable road. They go a step ahead from traditional SA design principles proposed by Endsley et al. (2003) which mainly relate to the physical and perceptual characteristics of a system, rather than how a system should function cognitively. These 16 principles are to:

1. Clearly define and specify SA requirements [of the system] [1].
2. Ensure role and requirements [of system agents] [1] are clearly defined.
3. Design to support compatible SA requirements.
4. Design to SA transactions [between human and non-human agents] [1].
5. Remove unwanted information.
6. Use customised or tailored interfaces.
7. Use multiple interlinked systems of roles and goals.
8. Consider the technology available and its impact on SA.
9. Ensure information presented to users is accurate at all times.
10. Provide appropriate and explicit communication links [feedback] [1].
11. Ensure that team members are cognisant of what other team members should know during the task performance [as in Intelligent Transport Systems] [1].
12. More information is not always better.
13. Use filtering functions.
14. Present SA related information in an appropriate fashion.
15. Use procedures to facilitate DSA [cognitive salient features][1].
16. Test DSA through the system lifecycle [drivability][1].

[1]*Added by the author*

The importance of system designers knowing what it is that different users of the system need to know during a task has been illustrated through in DSA in Chapter 2. A complex system design process should therefore begin with a clear definition and specification of SA requirements (design principle 1). This aspect of design is explained

in the context of driving task in the first half of this chapter through situation awareness requirement analysis. This is also a principle shared by SER design. SER that focuses on evoking correct driver behaviour (Theeuwes & Godthelp, 1995) from users by making a road homogenous, functional and predictable (Charlton et al., 2010). This matches with driver expectation and minimises confusion (Mackie et al., 2013).

Every road user interprets a scenario on the basis of his goals. A car driver can see a country road as a desirable challenging drive whereas for a learner car driver it can be just be a road to avoid fast-moving traffic around the city and dual carriageway. As a result, the car driver may adopt a higher speed, putting the learner at risk. This is not the case with motorways, where the road is very clear regarding what kinds of vehicles are prohibited such as agricultural vehicles, learner drivers, etc. and what it connects (usually inter-cities) thus giving a clear message to drivers that it is meant for high-speed driving to reach long distance destinations for business and other work (design principle 2).

Despite motorways being highly drivable their digital gantries often do not comply with design principle 3. Compatible SA means that instead of providing all information to everyone, it is worthwhile to present the information needed to pursue goals and fulfil the roles and SA requirements of different road users. For example, facts on heavy rain and winds at Inverness on a gantry on the M8 toward Edinburgh from Glasgow are useless for drivers commuting between Edinburgh and Glasgow. The thesis repeatedly emphasises the importance of SA transactions. An SA transaction is the exchange of SA related information to be integrated with the existing schema (Salmon et al., 2009). This principle (4) is often poorly supported in newer vehicle interfaces which flash warning lights for heavy fuel consumption when driving on rural B roads or roads with lower speed limits. In this circumstance where the communication link is poor between the road and the driver, the driver often chooses to speed. This problem can however be countered with design principles 5 and 6.

Design principle 7 is an area of future research. In a road system every agent has a different role and requirements but they are all interlinked at some level. Design principle 7 aims to offer road users separate but linked support systems. To do this, it is essential to determine how different road users interact with each other and with

different road characteristics (e.g. Salmon et al., 2014). This work is still at an exploratory level. The results need to be integrated into existing technology to propose a prototype interface or support system customised for different road users (trucks, cars, bikes, or motorbikes).

Design principle 8 relates to the very definition of human factor with the aim of "designing a system which is sensitive to human limitations and capabilities" (Proctor & Zandt, 2011, p.9. This is an obvious yet important concern for designers. Accurate information communication (design principle 9) is perhaps the backbone of maintaining the DSA of a system. For example if the sign of 'road works and reduce speed to 40 mph' is placed on the adjacent carriageway where no roadwork is carried out, instead of the carriageway where actually roadwork is happening, it will lead to traffic congestion, noncompliance with speed limits, and eventually accidents causing injuries.

Design principle 10 stems from the finding of Stanton et al. (2006) that links between different agents is far more important than actual DSA maintained by agents themselves. These links are maintained by feedback. This chapter partly throws light on this aspect; however, which form of feedback is important for which road user and how it can be supported through self-explaining design is an area for further research. This area has been explored by Walker (2002) in the context of vehicle design. Now the need of the hour is to extend this territory to study the feedback framework needed for user-centred road design.

Design principle 11 is a major concern for present-day road transportation. This is also supported by Salmon et al. (2014, p.193) who posit that "one road user is typically not aware of the other". In a road design context this is an area of system design intervention. In a user-centred road every agent should be aware of the other agents' specific views. This knowledge will guide the communication links (design principle 10) to be formed between different users of a system. For instance, the SER approach mainly focuses on the information held by the pedestrian and how this will affect car drivers. Design interventions such as traffic islands, pedestrian crossings, and speed cushions are used for efficient SA transactions between pedestrians and car drivers. Areas of incompatibilities need to be identified between other road users and customised communication links need to be established.

Design principle 12 (like design principle 10) takes into account the human limitation of finite information processing. More knowledge will not lead to more SA. Instead, more knowledge objects will lead to processing of irrelevant information and create a distorted view of the system which is not compatible with other agents. It will also lead to the very popular problem of increased workload. To prevent this design principle 13 recommends filtering the information to be presented to the road users. For example, in the in-vehicle map of lorry drivers it is far more important to indicate presence of bikes or pedestrians nearby so that they maintain a safe distance. This is due to the fact that larger vehicles can have a vacuum effect when driving at high speed, whereby driving too close to cyclists and pedestrians can pull them under the vehicle. This information is not necessary for car drivers however it is important for them to be aware of the traffic conditions ahead (congestion, road works, speed change, etc.).

It is not enough only to present information to support driver information needs. It is equally important to present them in a way that they are readily interpretable (Endsley, 2000). This is design principle 14 which is also the baseline of SER approach - that is, to design roads which explain themselves to users without any aid. This is facilitated by cognitively salient features (design principle 15). It is these features which are interacting actively with the road users, thereby maintaining the DSA of a system. Finally, design principle 16 recommends testing the DSA of a system. In other words, the aim of this principle is to determine to what extent a road is able to support SA transactions through compatible SA held by different agents. This is predictability, another core principle of SER. As put forth by Mackie et al. (2013, p.742), "predictability means that the look and feel of the road within each category should be very consistent so that the user behaviour is consistently reinforced". Feedback also contributes towards predictability. It is through feedback driver perceives the current status of the vehicles and what it is likely to do further (Walker, 2002). This thesis operates within these last three design principles (14, 15, and 16) to lay the foundation of a user centred road which can demonstrate all the above guidelines.

Cognitive scientists such as Chase and Simon (1973) and Gobet (1998) reported that a good model is sparing with its information. Less information will lead to more information being used from working memory, and that information is subjected to higher levels of perception. This is also echoed by Charlton et al. (2010). More stimuli

will not lead to a better self-explaining road (Charlton et al., 2010). The operator will only perceive stimuli which he perceives as important and use it to make further decisions. Keeping this in mind it is important to know where drivers are extracting information from. This knowledge will inform better design, as that particular source of information can be under the focus of transport engineers and other authorities.

4.3 Methods for Extracting SA Requirements

Situation Awareness Requirement Analysis was itself given by Endsley (1993) in the context of air-to-air combat for fighter pilots. It was proposed to discover the information that a system must provide to its users so that they can adapt as situations evolve (Groner, 2009). Requirement of SA for a task is determined through goal-directed task analysis (GDTA). There is some controversy here because GDTA is similar to the much longer established Hierarchical Task Analysis (HTA) in use since the 1970s (and not referenced in the former's development). The GDTA method follows Endlsey's definition of SA. GDTA starts with identifying the SA elements of a system. The second step is to put these elements within the appropriate level in Endsley's model. This second step is however only limited for dynamic systems with dynamic actors. The biggest advantage of defining SA requirement through GDTA is that it is technology-free and easy to administer (Endsley, 1993). SA requirement analysis has been successful in designing emergency plans for fire-fighters (Groner, 2009), for air traffic controllers (Endsley, 2000) and in advanced bomber missions (Endsley, 1993).

GDTA to identify SA Requirement (Endsley et al., 1993) has been developed through three steps: unstructured interviews, goal-directed task analysis, and structured questionnaires. It has undergone some modification since its inception by Groner (2009; 2012) in order to increase its sensitivity to complex systems and automated tasks as shown in Figure 4.1. In stage 1 the interviewees were asked a variety of questions such as, 'what is good SA', 'what would you want to know to have perfect SA', etc. Stage 2 sought to establish primary goals and goals adopted to achieve the former goals. The results of stage 1 and stage 2 were pooled to arrive at a list of elements, and stage 3 of structured questionnaires was administered. The results of these indicated towards the desire of operators to have information regarding higher level of SA available readily.

Endsley (1993) interpreted this as data which requires processing from lower levels such as tactical posture, survivability, etc. to provide to the operators with no effort.
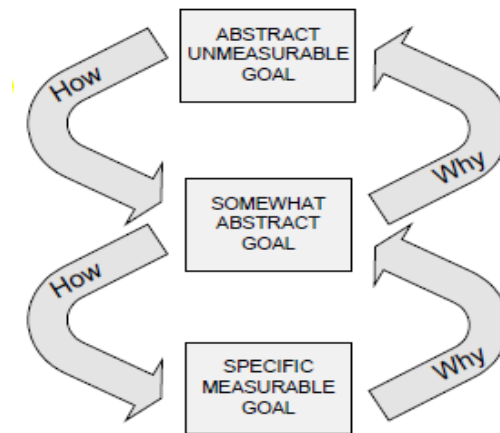
Figure 4.1. SA Requirement Analysis Process, adapted from Groner (2009)

Data collected from interviews is represented in the form of goal (Groner et al., 2011). Goals are viewed very differently by GDTA and a related task analysis method called Hierarchical Task Analysis (HTA). HTA was used by Annett & Duncan (1967) to understand complex tasks better (Stanton et al., 2013) to inform task and interface design. In GDTA, goals and decision questions are viewed as resources required by operators or unmanned vehicles to integrate into a system (Humphrey & Adams, 2010). However, HTA is more aligned towards a systems perspective as goals are understood as an end state of a task. GDTA is a derivative of HTA, but HTA has a major advantage over GDTA in terms of goal timing (Humphrey & Adams, 2010). HTA through plans and sub goals is very detailed. HTA can specify under what conditions and in which sequence a task needs to be completed to fulfil the requirements of subordinate tasks (Salmon et al., 2010). Furthermore, the origins of HTA date back to early 1900s and it has been used in various domains for over 40 years with no signs of major disagreement in human factor circles (Salmon et al., 2010). GDTA on the other hand has been used by Endsley et al. (2013) to determine SA requirement for pilots. Apart from this it has been used by Humphrey and Adams to determine SA requirement in a command and control paradigm (e.g. Adams, 2005; Humphrey, 2009; Humphrey & Adams, 2009). Hence, HTA has a much longer pedigree and serves the same conceptual purpose of task decomposition better.

In the preceding chapters it has been mentioned time and again that SA is an interaction-based phenomena. SA of a system is maintained through SA transaction and every agent in the system has compatible SA (Stanton et al., 2013) to do that. SA requirement analysis remains a useful exercise under this new paradigm. The SA requirement of the driver is thus defined by the quality of these interactions and how these transactions take place within a system. Hence, SA requirements in this experiment are measured by feedback and sources of information. Feedback is an important concept in driving (Walker et al., 2008). Norman (1987, p.29) has referred to feedback as, "sending back to the user the information what action has actually been done or what results have been accomplished." Depending on the quality of feedback the information can also contain the knowledge of results (Annett & Kay, 2005) if not actual results. Drivers and other road users are found to be very sensitive to feedback (Walker et al., 2008) irrespective of the modality (visual, auditory, tactile) (Horsewill & Coster, 2002). The thesis is written with the aim of designing user-centred road design and informing the role of distributed situation awareness in car driving. In this regard Chapter 4 provides the engineers and technology developers a basis from which to develop a design which maximises car drivers' safety and comfort, rather than overloading them and thereby degrading their performance. Chapter 4 also specifies what exactly needs to be known by car drivers, how that information is used by car drivers and how this supports SA transaction.

4.4 Methodology

4.4.1   Design

An existing HTA of the driving task (HTAoD) is used to systematically extract information requirements in terms of source and content. For this analysis, sources of information are the individual, the car and the road environment. Feedback modes are auditory, tactile and visual. In this HTAoD, 1100 driving tasks are grouped into basic, core; specific, general and global driving tasks. They are selected out of 1600 driving tasks for this thesis. Having finished coding of all the driving tasks selected, the inter-rater reliability (IRR) of the coding scheme is established. Many tasks are encoded as a combination of one or more types of feedback. For example, 'put car in 2nd gear' is categorized as both visual and tactile feedback. The driver through experience and

practice changes the stick to the 2nd-gear position and glances at the stick to confirm his action. Inter-rater reliability is especially important when using themes (Walker, 2002). Themes to some extent have a shared meaning and the IRR coefficient represents how well this meaning is shared (Walker, 2002).

4.4.2 Participants

After coding the driving tasks, two independent everyday British drivers (m=27 years, S.D = 1.414) were used for establishing inter-rater reliability.

4.4.3 Materials

For this analysis HTAoD by Walker et al. (2015) was used. Most human factor methods undergo an initial HTA before starting the actual analysis (Stanton et al., 2013). Similarly, SA requirement analysis in this chapter had to first go through HTA instead of the three-stage interviews of subject matter experts (SMEs). HTA is based on the fact that a task is defined by its objectives and its finished product (Walker, 2002). However, goals are needed to be set, plans are to be formed and tasks are to be performed to reach the final product (Annett et al., 1971). The end result of the HTA is an exhaustive description of task activity. For example, the overall task goal of 'driving a car' is broken down into its constituent goals, such as 'changing lanes', 'maintain speed of 70 mph' etc. To achieve this goal, plans like 'maintain adequate headway', and 'maintain sufficient rear distance' are to be formed. Finally, tasks like 'checking rear mirror', 'indicating', and 'turn steering wheeling to go left' have to be performed. McKnight & Adams (1970) divided driving into 7 main tasks which were subdivided in 43 tasks, which were broken down to 1600 intrinsically basic tasks. 400 plans were executed throughout the entire HTA process. In the present case the aim is to determine the information need of the drivers in every stage of task enactment. As SA requirement analysis is done to identify what is that comprises SA in the scenario under analysis (Stanton et al., 2013) which in this case is the physical act of driving which also happens to be a complex cognitive activity.

Other materials used are Microsoft excel for coding and predictive analysis software (PASW formerly called SPSS) for calculating reliability coefficients. HTAoD as given by Walker et al.,(2015) is made of task analysis conducted in 1970 by McKnight and

Adams, The latest edition of the UK Highway Code (based on the Road Traffic Act 1991), UK Driving Standards Agency information and materials, Coyne's (2007) Roadcraft (the Police/Institute of Advanced Motorists drivers manual), SME input (such as Police drivers) originating from research of Salmon et al., 2013, Salmon et al., 2010, Young et al., 2013 , etc. work in advanced driver training and numerous on-road observation studies involving a broad cross-section of 'normal' drivers. These sources are consistent with the requirement analysis.

4.4.4 Procedure

The HTAoD used in this analysis models the tasks as it should normally occur in Britain, thus the choice of this document is matched to this thesis. The next stage is encoding those tasks on the basis of themes. A theme-based analysis centres on encoding the meaning of phrases and sentences into shorter thematic units or segments (Weber, 1990). For example, for the phrase 'view road ahead next junction', the thematic segment might be as follows: the driver is deriving the information about the *road* through *visual feedback* from the *road*. HTA as given by Annett & Duncan (1967) is based on the idea that a task is defined by its end product and goals. Plans are needed to achieve those goals (Annett et al., 1971). Under this taxonomy 'driving a car' is broken down into subordinate tasks on the basis of goals and objectives. Goals are unobservable tasks goals of the driver (Stanton & Young, 1999). Tasks on the other hand are observable actions (such as 'check rear view mirror') which are performed by the car driver in pursuit of goals. Plans are also unobservable and they specify the sequence of task enactment needed, often dependent upon the necessary conditions being met in order to achieve the desired task goal. Since necessary conditions are required to be met to perform the next task, feedback is an essential feature in task analysis (Walker, 2002). A part of this chapter will also throw light on different modalities of feedback at different levels of driving.
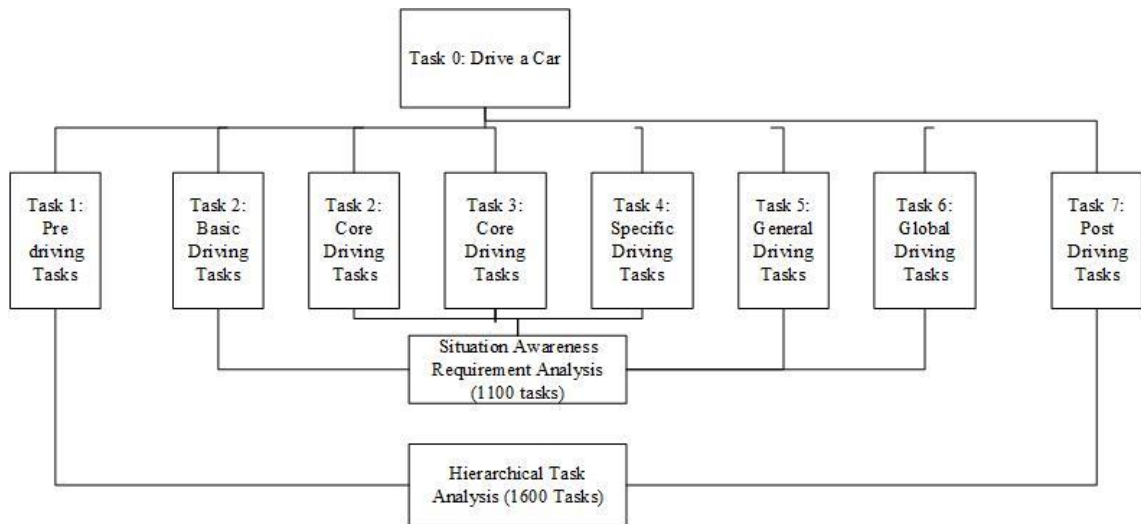
Figure 4.2. Top Level of HTA, adapted from Walker (2002)

The pre- and post-driving tasks are excluded in this experiment. This is due to the fact that the vehicle is stationary in pre- and post-task phases and there is no interaction with the road or other traffic. Hence, 1100 tasks obtained from 5 parent tasks were analysed for SA (Figure 4.2). Each of these tasks was coded in Microsoft Excel as 'I' (individual), 'C' (car), 'R' (road environment), 'V' (visual), 'A' (auditory), and 'Ta' (tactile). The driving context is a combination of driver, vehicle and road (Michon, 1993). As such this information is classified here under car, driver and road. Refer to Figure 4.3. This information classification also enables the identification of major agents in the environment which are interacting with the driver more than others.

Figure 4.3 SA Requirement Analysis of Driving (Microsoft Excel View)

4.5 Results

The HTAoD is used by three independent raters to systematically define the information needs of drivers in terms of source and modality. The independent raters make use of the same categorisation scheme as the experimenter. Across the six coding categories, inter-rater reliability was established at rho=0.943 for IRR1 and rho=0.921 for IRR2, both these values being significantly correlated (p>0.05 for IRR1 and IRR2, n=6). Having established satisfactory inter-rater reliability by two independent raters, results can now be analysed.

4.5.1 High-level Results

Out of the 1100 driving tasks analysed, 145 are basic driving tasks (fundamental physical control tasks of the vehicles performed by the driver); 139 are core driving tasks (tasks at this level are still very basic in nature and mainly deal with controlling vehicle in terms of speed and trajectory); 270 are specific driving tasks (tasks which deal with strategizing and making plans to adapt to environment-specific situations such as junctions and crossings); 280 are general driving tasks (tasks are increasingly goal-orientated and principal tactics for adapting to different road types - this also involves general procedures for negotiating with other traffic), and; 266 are global driving tasks (tasks to do with rule compliance and navigation, mainly relating to high-level driving strategies).

An interesting observation to be made here is that the lower-level tasks are fewer compared to higher-level counterparts. Fewer tasks mean fewer plans formed and executed. Specific and general driving tasks are the richest in task number. Specific and general driving tasks echo Endlsey's Level 2 SA: comprehension of the three-level SA model. A further observation reveals that the first 30 tasks of general driving are concerned with updating the driver's knowledge of the road, for example: 'observe movement of pedestrians', 'observe for slow moving agricultural machinery', etc. This means that these tasks are more cognitively demanding and require higher SA to negotiate the route successfully. Hence, a user-centred road should fulfil this SA need of the driver through appropriate SA transactions. In SER terms, this means special attention to SA should be given when changing road category e.g. when moving from dual carriageway to motorway, adjusting safety margins to accommodate traffic.

4.5.2 Basic Driving Task

Basic driving tasks are made of 'pulling away from standstill', 'steering', 'decrease speed' and 'speed control'. These tasks are fairly basic and occur at initial phases of driving. Significant difference is detected between visual and auditory feedback (t=2.15, df =144, p>0.05). Significant difference is also detected between tactile and auditory feedback (t=1.68, df=144, p>0.01). No significant difference is found between visual and tactile feedback. Similarly, significant difference is found in sources of information. Statistically significant difference is found among car and other traffic (t=1.45. df=144, p>0.05) and car and road (t=1.53, df=144, p>0.05) whereas differences between other traffic and road are non-significant. Figures 4.4 and 4.5 show the graphical representation of the results.
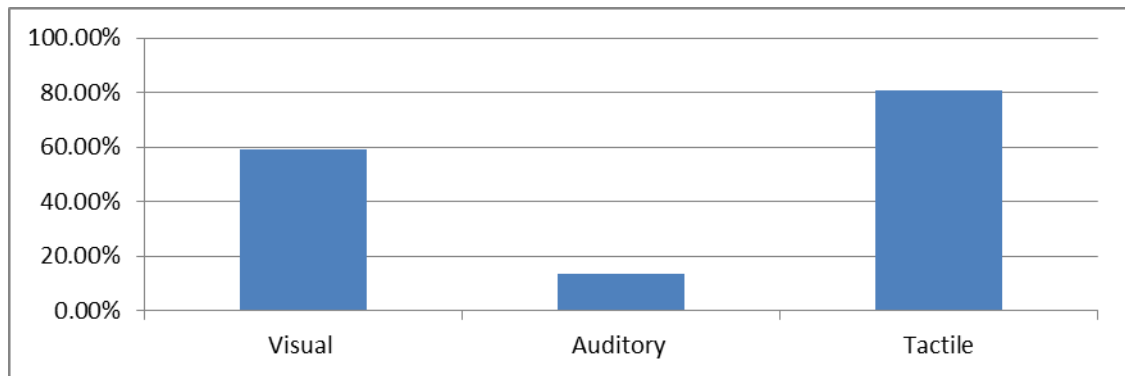
Figure 4.4. Basic Driving Task and Feedback
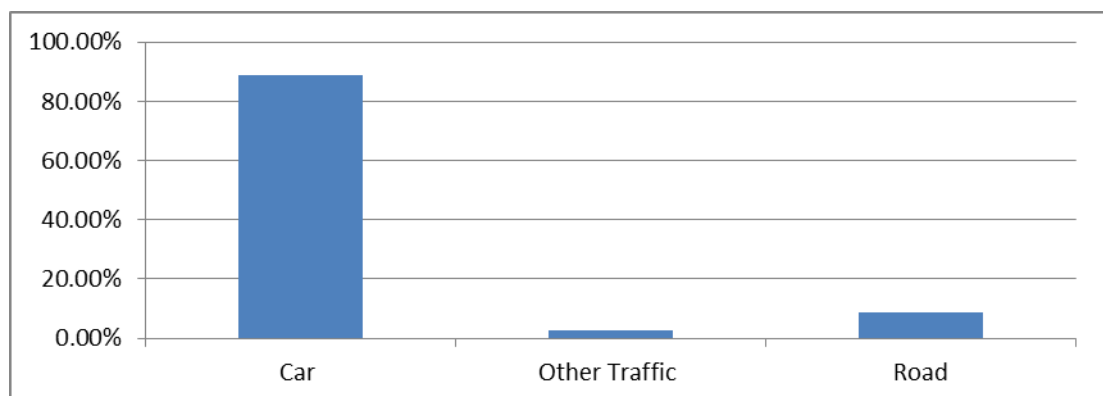
Figure 4.5. Sources of Information for Basic Driving Tasks

In light of the nature of basic driving tasks these results point towards a strong interaction between the driver and the vehicle. There is a clear vehicle dominance. Driver and car form a complex feedback system (MacAdams, 2003); the behaviour of the car results in certain behaviours of the driver, and vice versa. Drivers are fed information from the vehicle (e.g. from the speedometer, gear position, engine sound, mirrors) through visual and haptic means. Driving has more SA transactions taking place from agents with compatible SA within the vehicle. This is because drivers at the start spend considerable attentional resources to adjust to the vehicle. After this, drivers negotiate the route by activating the appropriate schema, and then alternate schemas and feedback as they go along the route.

4.5.3 Core Driving Task

Core driving tasks include negotiating bends and gradients, reversing, and exercising directional control. Statistically significant differences are found in all three pairs of feedback. Pair 1: Visual and auditory (t= 0.865, df=138, p>0.05), pair 2: auditory and tactile (t=0.814, df= 138, p>0.05) and pair 3: visual and tactile (t=0.822, df =138, p>0.05). For sources of information statistically significant differences is found between other traffic and road (t=0.75, df= 138, p>0.05). These tasks are a step above basic driving tasks in terms of complexity. In core driving tasks, drivers are required to assess the 'road surface classification', 'switch on interior heater' on climbing uphill, 'braking', and many others which require a combination of visual and tactile feedback.

An interesting observation to be made here is that core driving tasks also include reversing, but reversing is not only reliant on visual feedback as one might assume. Reversing includes 'depress brake', 'pull handbrake', 'depress clutch', 'turn upper body to face left of vehicle' – all sub-tasks which require constant tactile feedback from the vehicle. Another interesting point to be made here is that most of the preparation tasks (such as 'check mirrors', 'look over right shoulder' etc.) are very visually dependant; however those occurring when the car is actually in motion are mainly a mix of tactile and visual feedback. Furthermore, auditory feedback seemed to be most important through vehicle. This has been also confirmed by Walker (2002).
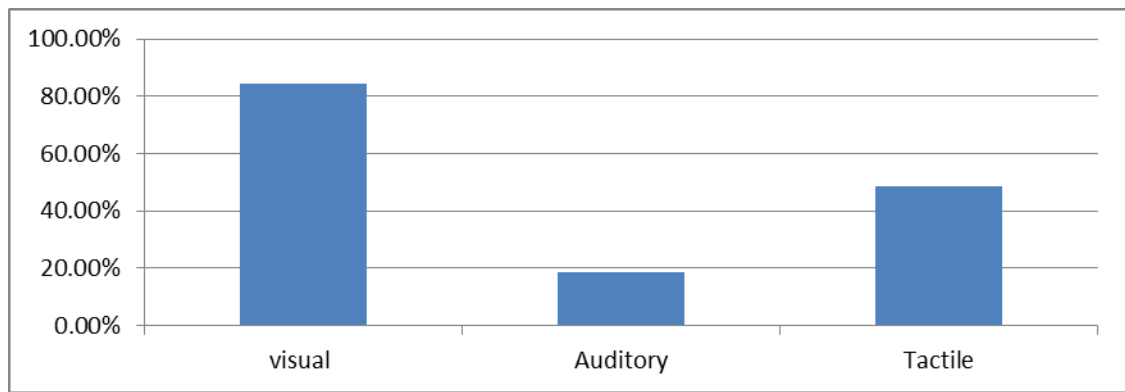
Figure 4.6. Feedback of Core Driving Tasks
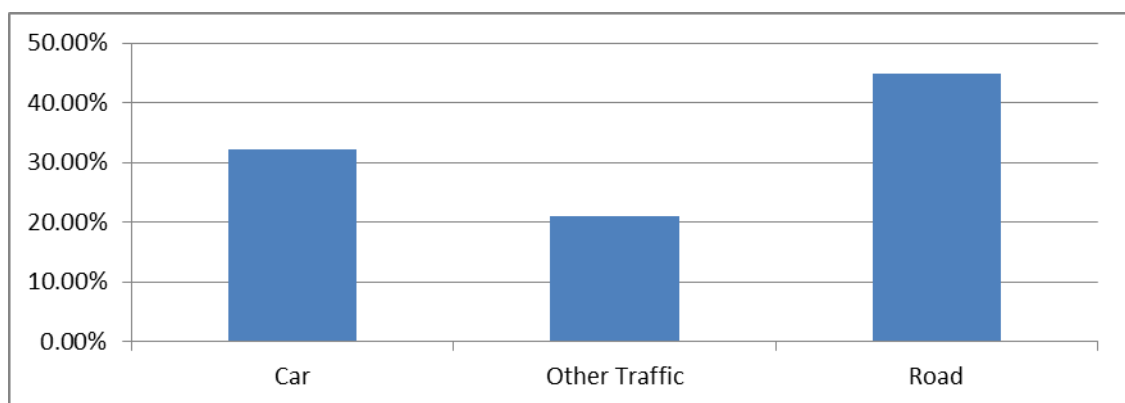
Figure 4.7. Sources of Information for Core Driving Tasks

Drivers are found to be deriving most of the information they require from the road in core driving tasks (Figure 4.7). A careful examination reveals that when the driver is on road the SA transactions are far more complex than that of basic driving tasks where driver is slowly coming to motion. This ties in well with the decision of excluding pre- and post-driving tasks from this analysis. All of the sources of information are worthwhile and SA transactions are taking place from car, road and other traffic. This further confirms the adopted systems perspective which views SA as not limited to drivers but as a property embedded in the system. From Figures 4.6 and 4.7 it can be interpreted that in core driving tasks (such as directional control, reversing, negotiating bends and gradients) SA transactions are very active from the road, and in the form of visual feedback. However, this changes when the core driving task is analysed in more detail. For example, while reversing tactile and visual feedback are found to be very important and this is fed through the car and the road equally. This indicates segments of road which experience large numbers of cars reversing should be designed to interact

smoothly with the vehicle and should be able to provide visual and tactile feedback to the driver.

4.5.4 Specific Driving Tasks

Specific driving tasks include 'emerging into traffic from other side', 'following other vehicles', 'overtaking', 'approaching junctions', 'deal with crossings' and 'leave junctions'. Statistically significant difference is found between visual and auditory ($t=2.23$, $df =269$, $p>0.01$) and visual and tactile ($t=2.14$, $df =269$, $p>0.05$) feedback. As a source of information road is found to be statistically different from car ($t=2.11$, $df =269$, $p>0.01$). From Figure 4.8 it is very clear that all specific driving tasks require visual feedback. Additionally, one third of these tasks require both tactile and visual feedback, and almost 17% require a combination of visual, auditory and tactile feedback. As mentioned previously, specific driving tasks include strategizing and making plans to negotiate immediate, challenging road environments, thus these are cognitively more demanding than the above two (basic and core) task types.
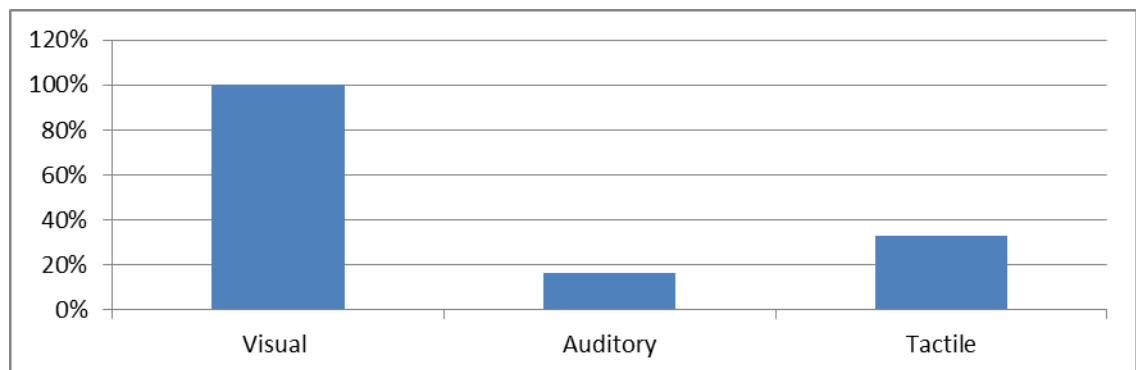
Figure 4.8. Feedback in Specific Driving Task
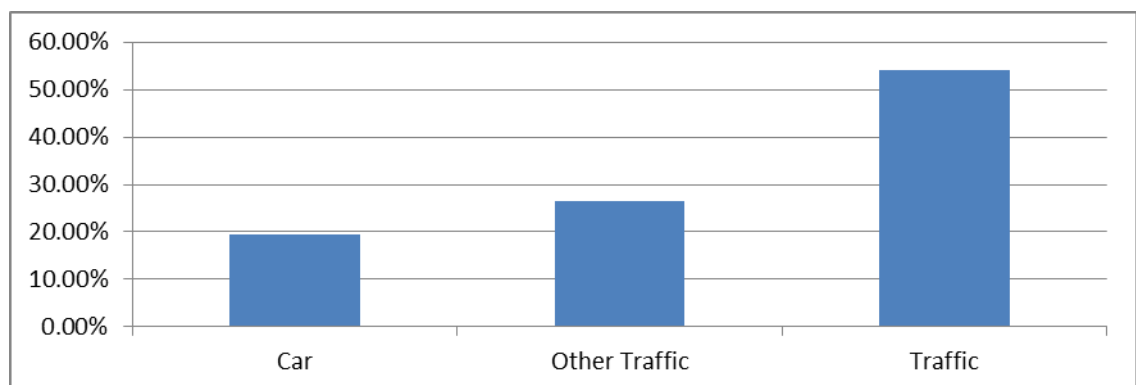
Figure 4.9. Feedback and Sources of Information for Specific Driving Tasks

In this task type drivers have the highest interaction with road (Figure 4.9), thus these results need to be interpreted with caution for this thesis. The first caution to exercise here is to avoid overemphasizing or over-representing visual feedback. Walker (2002) reported that visual feedback in specific driving tasks misled SA probes, and led to higher false alarms. Walker (2002) further reported that the highest workload in drivers occurred when visual feedback is used alone.

As put forth by Hendy (1995, p.6), 'good SA is reported to be associated with rapid goal achievement through timely and appropriate actions in response to sensory input'. Adding feedback will mean improving the SA of the system. In terms of design it is wise not to overload drivers with visual information in the form of road signs and signboards. Instead, a few well-timed pieces of visual information can instigate safe driver behaviour. This was confirmed by Wallis et al. (2007) where experienced drivers were found to be successfully manoeuvring their steering without any knowledge of steering wheel angle or where the vehicle was heading. They could do this only with the help of well-timed samples of visual information (Wallis et al., 2007). Vision is the most important requirement in driving and therefore visual feedback is of utmost importance (MacAdams, 2003). The exact level of visual feedback varies depending on the task under analysis (MacAdams, 2003). But other modalities are too important to be ignored while considering design of a system, as they have the ability to make a positive impact to design. Such implicit feedback has already had positive impacts in vehicle design (Walker et al., 2008). Visual feedback will clearly help the driver to drive a vehicle satisfactorily, but user-centeredness is much more than mere satisfaction (Walker, 2002).

The importance of multimodal feedback can be illustrated with an example. Under the specific driving task of 'dealing with crossings', the driver demands 35% tactile feedback, 30% auditory feedback and 100% visual feedback. In Kerang crash the truck driver didn't receive any auditory feedback and relied completely on visual feedback from the RLX (Salmon et al, 2013). This led to the LBFTS error leading to a crash with the passenger train. Furthermore, drivers need highest information from road (65%), other traffic (30%) and vehicle (5%).

If a road is designed with no consideration for auditory and tactile feedback, the driver will have poor SA, and because many SA agents who have compatible SA will be unable to contribute to SA transaction, this will lead to poor SA of a system. In this case, the driver and pedestrians have to increase their SA to compensate for this SA loss, and eventually the crossing or any place where there is a likely interaction between vehicle and pedestrian will be deemed dangerous. However, with SER these results can be taken under consideration and such risks can be mitigated. For example, it has been found that if a smooth surface succeeds a rough surface there is a mean reduction of 5% speed (TeVelde, 1985), causing the driver to automatically slow down near the crossing without expending much mental effort.

4.5.5 General Driving Tasks

General driving tasks are very similar to specific driving tasks. The difference lies in the generalizability of the plans and strategies to deal with different roads types. Specific driving tasks serve to negotiate immediate road features, and have strategies and operations to deal with junctions, crossings, etc. (Hanson, 2001). General driving tasks include increasingly goal-oriented and overarching strategies to adapt to different road types or general procedures for negotiating other traffic (Hanson, 2001). General driving tasks include four major tasks: 'deal with different road types', 'deal with different road hazards', 'general tasks for reacting to other traffic' and 'emergency manoeuvres'.

Statistically significant difference is found between visual and auditory feedback (t= 1.45, df=279 p>0.05) and visual and tactile feedback (t= 1.75, df=279, p>0.00). Statistically significant difference is found between 'car and road' (t=1.26, df=279, p>0.00) and 'car and other traffic' (t=1.21, df=279 p>0.05) as sources of information. Curiously, other traffic and road are not statistically different from each other (Figure 4.11). This means both are important sources of information for the driver, and agents in both sources are almost equally active in SA transactions.
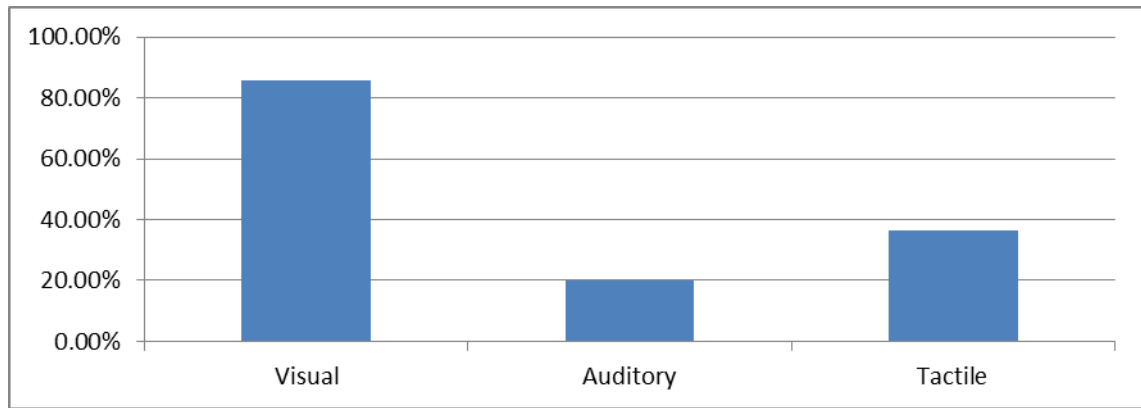
Figure 4.10. Feedback in General Driving Tasks
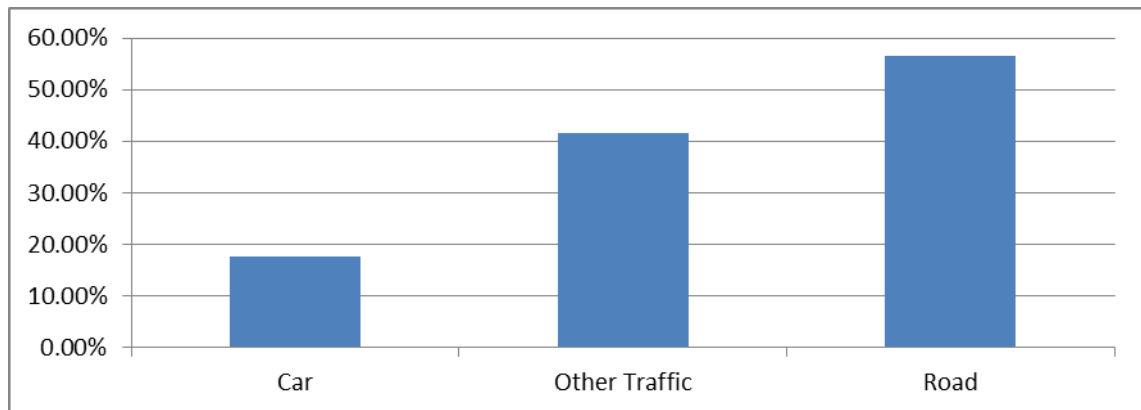
Figure 4.11. Feedback and Sources of Information for General Driving Tasks

These results follow the previous trend of drivers' need for multimodal feedback (Figure 4.10), and the importance of good design of road and vehicle. Interesting results can be seen in 'emergency manoeuvres' of general driving tasks. Drivers demand 65% of tactile feedback, 53% of visual feedback and 13.25 % of auditory feedback from road (37%), car (41%) and other traffic (21%). This means in an emergency situation a driver's SA will be ideally made of equal information from road and car. In other words in an emergency driver would preferably be in a *forgiving road*; a concept related to SER. Forgiving roads constitute road environments which contain basic tools for mitigating accidents caused due to driver behaviour, which constitute 80% of total accidents (Bekiaris et al., 2011). In 'general tasks reacting to traffic' drivers have indicated that they would prefer visual feedback from other traffic, with little feedback from other modalities. This answers a long-standing question of why two-wheelers such as cyclists and motorcycles are potentially dangerous. Due to their different size they are often not visible (Mundeteguy, 2011) or are obscured by other car drivers. Exploring this interaction in detail is beyond the scope of this thesis.

4.5.6 Global Driving Tasks

Global driving tasks relate to 'surveillance', 'navigation', 'rule compliance', 'task related to environment', 'IAM system of car control', 'vehicle/mechanical sympathy', 'driver attitude/deportment'. Significant difference is found between visual and auditory feedback (t= 2.38, df=265, p>0.05). No statistically significant difference is found among sources of information and in any other modalities of feedback. Global driving tasks relate to high-level driving strategies and are mainly concerned with rule compliance and norm driving. The fact that none of the sources of information are statistically different from each other means that the interaction is highest at the global driving task level.
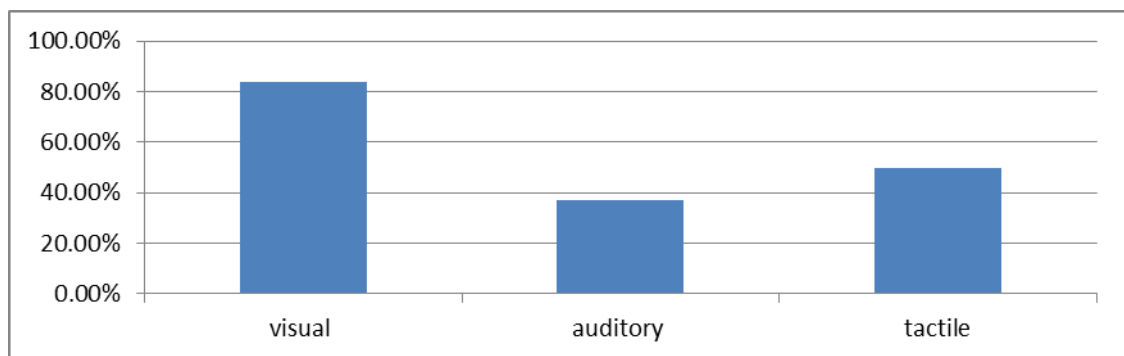
Figure 4.12. Feedback in Global Driving Tasks
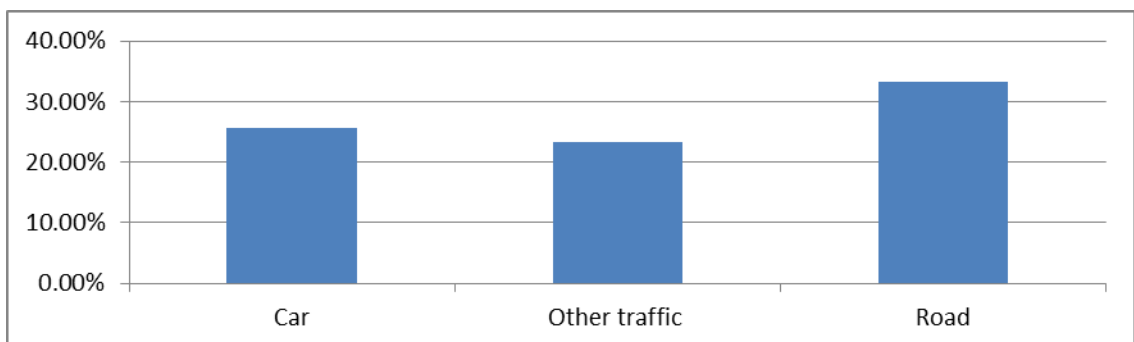
Figure 4.13. Feedback and Sources of Information in Global Driving Tasks

Interesting results can be seen in 'rule compliance'. Drivers derived 100% information from the road but rely on visual, tactile and auditory feedback for each and every rule compliance task from the road. Clearly, this suggests that drivers are demanding for a SA-centred SER or a user-centred road where none of their senses are overloaded. One

of the two tasks under rule compliance is to 'act on guidance/instructions/rules given by Highway Code'. This does not mean that the driver needs to be subjected to endless training programs until he/she complies with the rules in the code. Instead it means a road should be purposefully designed such that the driver automatically follows the rule. If drivers are found to be violating speed limits in an urban area repeatedly then no amount of negative feedback (e.g. fines or driving classes mandated by the authorities) is going to correct it. These measures only work if drivers are knowingly driving faster than they should (Bekiaris et al., 2011), but if they do not, then several traffic calming measures such as speed bumps, road narrowing, or fishbone design can help in decrease speed. This can be again achieved by a user-centred road design which is the aim of this thesis.

4.6 The Role of Road Design

Figure 4.14 shows the road as the main source of driver SA. Drivers are found to be deriving 48% of SA exclusively from the road in this analysis of 1100 driving tasks. Walker et al. (2013) has confirmed this through a naturalistic study. Walker et al. (2013) reported that 42% of information contributing to driver SA was found to lie externally to the driver himself. Therefore these results confirm that driver SA lies in the environment external to himself (Figure 4.14).
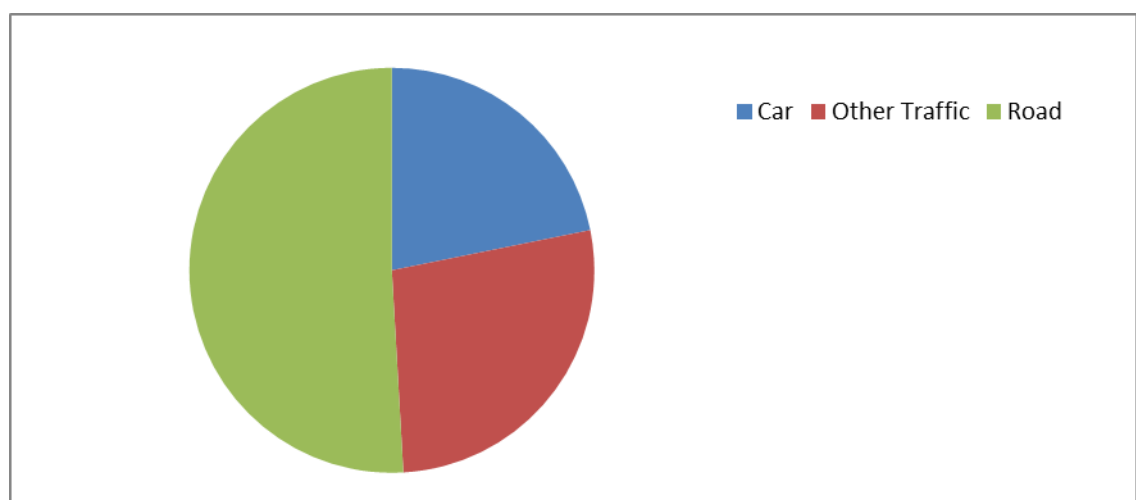
Figure 4.14. Road as the Main Source of SA

Basic driving tasks which constitute vehicles coming to motion and merging into the main road have reported 'vehicle' to be the main source of information. In core, specific and general driving tasks, when the vehicle is on the road, in motion, and is surrounded by other traffic, 'road' is the main source of information. This proves two points. One of these is that choosing DSA from a literature of disparate definitions and models has been the correct choice. Second is that the secret of accident reduction and driver safety lies with the design of the road infrastructure and environment. The current problem is that road which is in fact a socio-technical system is studied in isolation. Drivers are picked and trained over and over in addition to attending driving classes for a violation. The immediate need is to study the road as one system of which drivers are a part like any other infrastructure features. This calls for a paradigm shift from training, educational programs, assistance vehicle systems to user-friendly road design.
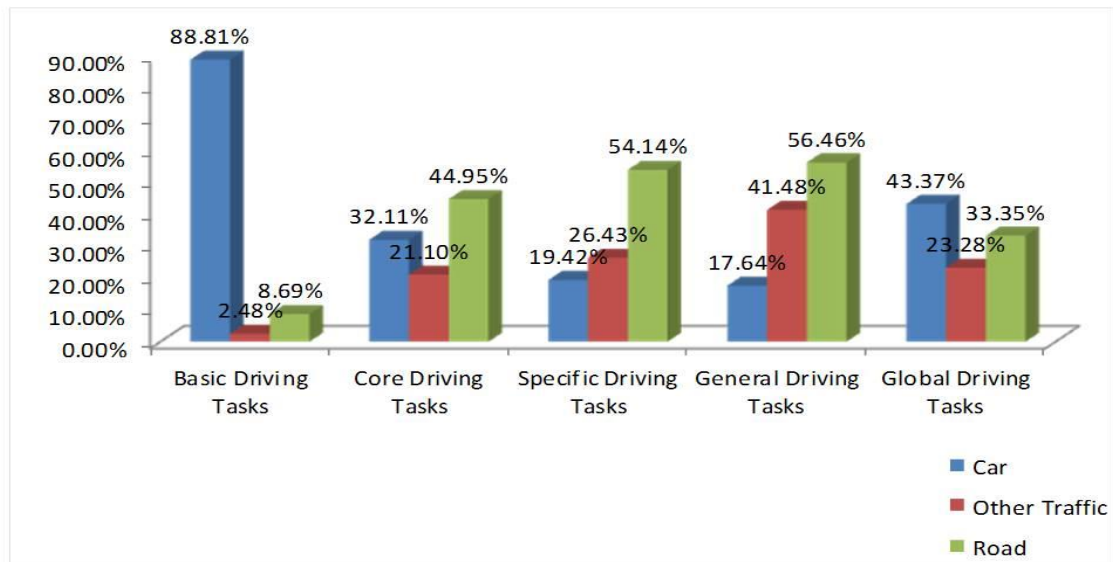
Figure 4.15. Driver Information Needs from Different Sources of Information

In driving task analysis, behaviour moves from basic low-level physical tasks to high-level cognitive tasks (Hanson, 2003). This evolution of tasks occurs through plans, operation, rules, and procedures, finishing up and overarching driving strategies (Hanson, 2003). Figure 4.15 shows how the role of roads, vehicles and drivers differ with different levels of driving. It can be seen that the role of information source in driver SA differs as the driving task increases in hierarchy. Global driving tasks which include high-level tasks of 'surveillance', 'navigation', 'rule compliance' and many

more require almost equal input from car, vehicle and road. Global driving tasks did not report any significant difference between sources of information. Changing the profile of information needs, along with communication links (feedback), at different levels of driving tasks has design implications. For instance, in the case of accidents involving cars at a clearly viewable junction on a particular road, the attention of transport engineers should focus on the design of that junction in terms of signs, gradients involved, number of roads and slips. This is due to the fact that during specific driving tasks while negotiating a junction, a driver's main source of information lies in the road environment. This is also one of the benefits of doing a requirement analysis.

Visual stimuli are often overrated in driving. It has been reported that 90% of information processed visually is lost during higher-order processing (Gregory, 1970). In light of this it is essential to provide information to drivers through haptic, tactile and auditory means. A secondary result from this SA requirement analysis is the evidence for the need for multimodal feedback at all levels of driving. The results of this analysis suggest that visual feedback is of highest significance in driving. Although visual information is dominant, most of the information is perceived through a combination of either auditory, tactile or both means. Auditory feedback is found to be found to have the most effect when implemented as a supplement (MacAdams, 2003). Complex systems perform better in a multi-feedback framework. Therefore practitioners should not overlook auditory and tactile feedback while designing a system.

Vehicles with high amounts of feedback have already reported an increase in driver SA (Walker et al., 2008). DSA has proposed to develop explicit communication links (design principle 10) amongst agents in a system. Detailed analysis of specific driving tasks show that tasks related to following other vehicles and merging into traffic do not have a specific need for tactile or auditory feedback. Although tasks involving interaction with road features such as 'approach junctions', 'deal with junction', 'deal with crossings' and 'leave junctions' call for more than just visual feedback. Thus to design a user-centred junction in a user-centred road, it is recommended that the on-junction exit is designed with a slight change in the gradient (undulation), indicating the change of road and allowing the genotype schema to be activated.

4.7 Study 2: Driver SA in Practice

Study 1 tells us about the total information need of the drivers. What it does not tell us is how those information needs become manifest when placed in contact with a real road environment. This is the purpose of study 2. It seeks to provide insight into the following question: out of all the possible features identified in study 1, which ones are the more or less 'cognitively salient'? This question lies in the heart of this thesis and of an objective to link between SA and SER. The identification of these features in SER literature has proceeded along various lines and a good summary is provided by Charlton et al. (2010). Examples include several recent papers which use a form of 'picture sorting' task to discover what road features distinguish different road types (e.g. Weller et al., 2008; Stelling-Konczak et al., 2011). Other methods include questionnaires (e.g. Goldenbeld & Van Schagen, 2007) and driving simulator studies (e.g. Aarts & Davidse, 2007). Curiously, very few studies take place in real road environments, and no studies have been identified which make meaningful reference to, or use of, SA as a concept. Study 2 is a pilot study undertaken to explore these linkages and place the findings of study 1 in contact with real-world driving.

4.8 Methodology

4.8.1 Design

This is a pilot study and before significant resource investment in primary data collection is embarked upon (Chapter 5), an initial data set is analysed. The original data was collected by Walker (2002) for investigating the effects of feedback on vehicle design. The work represented in this study extends this original work by creating various forms of additional analysis such as semantic networks and social network analysis to provide further insights into cognitively salient features. The experience gained in performing this analysis informed the large-scale naturalistic study discussed in Chapters 6 and 7 of this thesis.

Study 2 uses semantic networks to investigate the source of driver SA in a real world study. Semantic networks are an established way of representing knowledge (e.g. Collins & Loftus, 1975; Collins & Quillian, 1972). Creating networks to represent driver SA involves extracting the noun-like information elements from drivers (called

nodes or vertices in semantic networks) and establishing verb-like links (or edges) between them. The result is that when elements become temporally, spatially, causally or semantically interlinked they begin to form 'concepts'. It is claimed that one can produce dictionary-like definitions of concepts and that a definition of any situation (or state of knowledge) can be represented (Oden, 1987). Networks capitalise on a number of important principles of systems thinking. The first is the Aristotelian idea that "the whole is characterized not only by its parts, but by the relations [or mapping] between the parts". SA can, therefore, exhibit the synergistic property of being more than the sum of its information elements (thus responding to the issue above concerning normative forms of SA). Networks also represent the individualistic nature of awareness, and the propagation of activity through them need not be linear. Sticking to the tractable nature of SA as concept is the fact that methods such as concurrent verbal protocols) exist to extract information from the working memory of drivers, and methods also exist to construct networks based on that data (e.g. Ogden, 1987; Smith, 2003). In the remainder of this thesis these methods will be deployed in order to provide a theoretically robust way of exploring and extracting cognitively salient, endemic road features from real-world settings. The verbal protocol data was then subject to a simple encoding scheme, one that coded phrases non-exclusively into four themes: 'own behaviour', 'behaviour of the vehicle', 'road environment' and 'other traffic'. Two different analysts encoding the same transcript achieved significant correlations at the 5% level (p = 0.7 for analyst 1 (n = 756) and p = 0.9 for analyst 2 (n = 968) as did the same analyst re-encoding the same transcript later in time (p = 0.95; n = 756; p < 0.01).

4.8.2 Participants

The results are drawn from a sample of 12 male drivers with an age range from 20 to over 50 years of age. Driving experience ranged from 3-44 years. Each driver possessed a valid UK driving licence with no major offences. Furthermore, each driver reported 10-12,000 miles of driving every year. As a control measure a 30-item DSQ questionnaire (West et al., 1992) was administered to participants. Mean score of the DSQ was 3.25 (S.D = 1.63). Participants who exceeded the mean score by ±1 were excluded from this experiment.

4.8.2   Materials

In the original experiment the participants drove a 14.5 mile long route in the West London area. An additional 3-mile long route was also used as warm up route. The route is comprised of one motorway section (70 mph speed limit for 2 miles), seven stretches of major road (50/60 mph speed limits for 6 miles), two stretches of rural road (60 mph speed limit for 3 miles), three stretches of urban roads (40 mph limit for 1 miles), one residential section (30 mph limit for 1 mile), and fifteen junctions (less than 30 mph speeds for 1 mile). Experimental runs took place at 10:30 in the morning and 2:30 in the afternoon (Monday–Thursday) as well as 10:30 on Friday. These times avoided peak traffic hours for the area, and all runs were completed in dry weather. Post-trial desktop applications such as Notepad, Leximancer™, Agna and Microsoft Office® were used.

4.8.4 Procedure

Formal ethical consent was obtained from all participants before the study commenced with particular emphasis placed on control of the vehicle and safety of other road users remaining the participants' responsibility at all times. An instruction sheet on how to perform a verbal commentary was read by the participant, which described that they should drive as they normally would but provide a constant commentary about why they were performing current actions, what information they were taking from the environment, how the vehicle was behaving and what actions they planned to take. Drivers were instructed to keep talking even if it appeared to them that what they were saying did not make obvious sense. The experimenter provided examples of the desired form and content that this should take.

There then followed a warm up phase. A three-mile approach to the start of the test route enabled the participants to be practised and advised on how to perform a suitable verbal commentary. This involved providing suggestions and guidance from the passenger seat and pulling over to review the audio transcript and advise where necessary. All participants were able to readily engage in this activity with a mean word per minute rate in excess of 30. Minimal advice was needed. The verbal commentary acted as a form of secondary task and the high rate of verbalisations seem to indicate spare capacity and little grounds to suspect interference with the primary driving task.

During the data collection phase the experimenter remained silent aside from offering route guidance and monitoring the audio capture process. The drivers, meanwhile, provided a constant verbal commentary as previously instructed. Drivers were debriefed upon return to the start location.

4.8.5 Data Analysis

The verbal commentary data was then subject to a simple encoding scheme, one that coded phrases non-exclusively into four themes: 'own behaviour', 'behaviour of the vehicle', 'road environment' and 'other traffic'. Two different analysts performed a blind encoding of each transcript and achieved significant correlations at the 5% level: Rho = 0.7 for analyst one and Rho = 0.9 for analyst two.

The data is then treated with Leximancer™ which automates the creation of semantic networks, and does so with complete repeatability. This is one of the two additional analyses performed in this thesis to align it with the aim of laying the groundwork for user-centred road design.

Leximancer™ uses text representations of natural language in order to create themes, concepts and links. This is achieved by algorithms which refer to an in-built thesaurus on the one hand, and to features of text such as word proximity, quantity and salience on the other. Leximancer™ has been used extensively in previous studies. It has provided insight into organisational change (Rooney et al., 2010), intergroup communication (Hewett et al., 2009), analysis of web content (Coombs, 2010) and large scale meta-analyses of themes within scientific journals (Cretchley et al., 2010). The application of this technique to a more intensive form of analysis based on concurrent verbal protocols of real-life transport contexts is a novel one.

Leximancer™ subjects this raw textual data to six main stages in order to create the semantic networks:

1. Conversion of raw text data (definition of sentence and paragraph boundaries, etc.).
2. Automatic concept identification (keyword extraction based on proximity, frequency, and other grammatical parameters).

3. Thesaurus learning (the extent to which collections of concepts travel 'together' through the text is quantified and clusters formed).

4. Concept location (blocks of text are tagged with the names of concepts which they may contain).

5. Mapping (a visual representation of the semantic network is produced showing how concepts link to each other.

6. Network analysis (this stage is not a part of the Leximancer™ ™ package but was carried out as an additional step to characterise the structural properties of the semantic networks).

4.9 Results

A total of six networks are derived from this process, one for each road type. The networks provide a representation of the knowledge extant in the driver's working memory as they encountered each environment. They have been produced automatically with complete repeatability from raw verbal transcripts: no manipulation of the data or the process of network creation was undertaken. In Leximancer™ the noun-like knowledge objects extracted from the verbal transcript are referred to as 'concepts' and each one is ascribed a relevance value from 0 to 100%. This value is derived from the number of times the concept occurs as a proportion of the most frequently occurring concept (Smith, 2003). In total, 174 concepts were extracted from the six semantic networks. In order to reduce the data to the highest scoring concepts and to avoid the inevitable idiosyncrasies of low scoring, highly personal and infrequently occurring concepts, those which scored lower than two standard deviations of the mean in each individual road-type category were excluded. This gives rise to a high-scoring subset of 25. These were used to create Table 4.1 in the next page.

Table 4.1 : List of Behavioural Features Crossed With Road Types

Concept	Example	Motorway	A/B Roads	Urban Roads	Country Roads	Residential Roads	Junctions
Ahead	"can't really see what's going on ahead"	■					
Behind	"that guy behind has given us some space"	■				■	
Bend	"and just accelerating off round the bend"				■		
Braking	"braking for junction"				■		
Car	"noticed hazard lights on car in front"	■		■		■	
Check	"just checking my blind spot over my right shoulder"		■				■
Clear	"it's all clear ahead"						■
Coming	"um, because there's no other cars coming"	■					
Corner	"got this corner here, looks a bit narrow"				■		
Doing	"speed limit here is 60, but we're only dong 40"		■				
Fourth	"into fourth, 30mph"		■				
Front	"the Volvo in front isn't gaining any road speed noticeably"	■					
Gear	"so we're in 5th gear cruising along at 45 now"				■		

Concept	Quote						
Hill	"accelerating more to get up hill"				■		
Indicating	"indicating left"						■
Lane	"as I observe the, er, stay in lane command, or whatever it is"	■	■				
Mirrors	"just occasionally checking the mirrors"						■
MPH	"now a 60mph limit"		■				
Parked	"want to keep a good distance from the parked cars"			■			
Pull	"but it pulls along quite nicely"		■	■		■	
Road	"road narrows at this point"	■		■			
Around	"not sure what the speed limit is around here"			■	■		
Slow	"so I'll slowdown"		■		■		
Take	"just gonna take it easy"				■		
Third	"keep it in third for a minute"				■		

What is clear is that marked changes occur in the content of driver SA as they enter and exit different road environments, to such an extent that no concepts are common across them; Table 4.2 highlights this point. It presents a summary of the limited extent of overlap across all combinations of road types, with a maximum value of only 11.5% (two concepts) occurring between motorways and urban roads. Road type is thus a powerful contingency factor in subsequent driver SA, a finding which accords with the varied literature on self-explaining roads.

Table 4.2: Percentage of shared concepts across different road types.

	Motorway	A/B Roads	Urban Roads	Country Roads	Residential Roads	Junctions
Motorways		3.8	11.5	0.0	7.7	0.0
A/B Roads	3.8		3.8	3.8	3.8	3.8
Urban Roads	11.5	3.8		3.8	7.7	0.0
Country Roads	0.0	3.8	3.8		0.0	0.0
Residential Roads	7.7	3.8	7.7	0.0		0.0
Junctions	0.0	3.8	0.0	0.0	0.0	
Totals	23.1	19.2	26.8	7.6	19.2	3.8

Analysis of these networks now proceeds on the basis of their structure and content. This is second additional analysis performed on this data, the first one being structural analysis using Leximancer™ The structural analysis employs techniques from graph theory to view the semantic networks in terms of nodes (n) and edges (e). These procedures help to reveal important underlying structural properties of the semantic networks which are not readily apparent from visual inspection alone. The metrics used are: density, diameter and centrality.

Density is given by the formula:

$$Network\ Density = \frac{2e}{n(n-1)}$$

Where e represents the number of edges or links in the semantic network and n is the number of nodes or semantic concepts. The value of network density ranges from 0 (no concepts connected to any other concepts) to 1 (every concept connected to every other concept; Kakimoto et al., 2006). Density is a metric which refers to the semantic network as a whole and is a measure of its overall level of interconnectivity. Higher levels of interconnectivity suggest a richer set of semantic links and a well-integrated set of concepts. A denser network is also likely to have better connected concepts and

shorter average path lengths. In order to diagnose the latter, a further metric is employed: diameter.

Diameter is given by the formula:

Diameter $= \max_{uy} d(n_i, n_j)$

where $d(n_i, n_j)$ is "the largest number of [concepts] which must be traversed in order to travel from one [concept] to another when paths which backtrack, detour, or loop are excluded from consideration" (Weisstein, 2008; Harary, 1994). Diameter, like density, is another metric which refers to the network as whole. Generally speaking, the bigger the diameter the more concepts within the semantic network that exist on a particular route through it. Again, generally speaking, a denser network will have smaller diameter (because the routes across the network are shorter and more direct) while a less dense network will have a larger diameter (as routes across the network have to traverse a number of intervening semantic concepts). This measure is related to the idea of clustering and to individual semantic concepts which are more or less well connected than other concepts. In order to diagnose this facet a further metric is deployed: centrality.

Centrality is given by the formula:

$$Centrality = \frac{\sum_{i=1;j=1}^{g} \delta ij}{\sum_{j=1}^{g} (\delta ij + \delta ji)}$$

Where g is the number of concepts in the semantic network (its size) and δji is the number of edges (e) on the shortest path between concepts i and j (or geodesic distance; Houghton et al., 2006). Centrality gives an indication of the prominence that each concept has within the semantic network. Concepts with high centrality have, on average, a short distance (measured in edges) to other concepts, are likely to be well clustered and to be near the centre of the network. Concepts with low centrality are likely to be on the periphery of the network and to be semantically distant from other concepts.

4.9.1   Semantic Networks

The twelve drivers contributed a total of 27,225 words into the analysis (Mean = 4537; SD = 2343). As Figure 4.16 shows, across all road types' drivers spent just over half of the time (54%) talking about factors external to themselves. Of this, 42% of the verbalizations refer explicitly to hazards, with verbalizations concerning the external road environment accounting for (18%) and, largest of all, other traffic accounting for (24%). Here we see an indication of the systemic, generative nature of the driver – vehicle – road system. The driver's own behaviour, and that of their vehicle, enables them to 'sample' the external road environment (including other traffic), which in turn 'modifies' their awareness of that environment, which in turn 'directs' further behaviour, and so on.
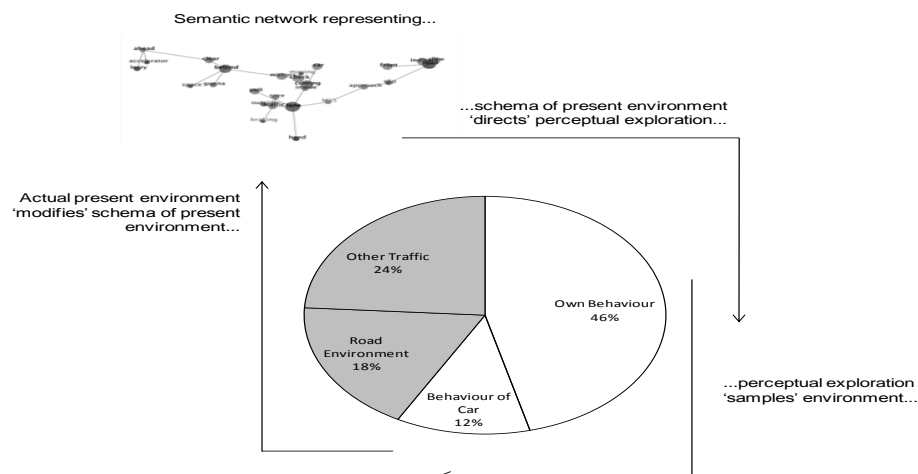
Figure 4.16. Role of the Road Environment in Driving

Figure 4.16 presents the results of applying density, diameter and centrality to the 12 semantic networks created from the verbal protocol data. The radar plots (shown in Figure 4.17) show that the structural metrics also have some contingency upon road type.
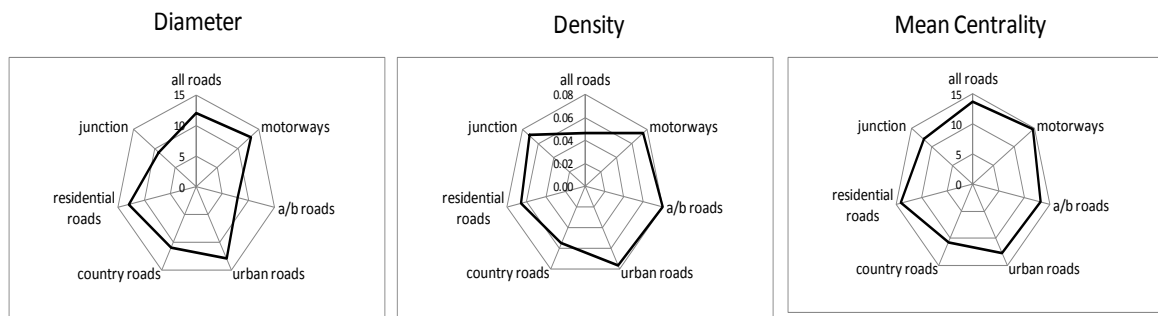
Figure 4.17. Radar Plots of Network Density, Diameter and Centrality and their Contingency upon Road Type

Motorway: Driver SA in this environment is dominated by concepts which relate to the road environment itself. The concept of 'lane' scored 100% and 'road' scored 73%. As discussed at length above, motorways represent a form of total environment and are regarded as an exemplar of a self-explaining road. Driver SA seems to mirror this in that the road environment itself seems key to the driver's awareness. The third-highest scoring concept is 'behind' (83%). This refers to events that are encroaching from this direction, something which the multi-lane configuration and high speeds of a motorway afford. The clustering of concepts is relatively high (i.e. mean centrality = 14.59), as is the distance across the network (i.e. diameter = 13), and overall the network is quite sparsely interconnected (i.e. density = 0.07). This would seem to suggest that a number of distinct key concepts are simultaneously active in driver's SA.

Major A/B Classification Roads: Driver SA in this environment is dominated by one particularly high scoring concept: 'slow' (scoring 100%). Relatively speaking, network density is high (0.08) and diameter low (8). Here an individual concept links to most other critical concepts, which seems evident in the other high scoring concept called 'doing' (71%), as in "I'm doing 60 mph". Clearly the awareness of speed, and reductions thereof, colours many other aspects of SA in this setting.

Urban Roads: Driver SA in this environment seems to be characterized by other cars (100%) and their interactions. Indeed, a number of other concepts hint at new goals, such as 'parked' [car] (48%) and 'pull' (as in 'pulling away': 70%). Interestingly, the wider environment features in the concept of 'around' (48%), e.g. "not sure what the speed limit is around here" or "being a bit cautious around here". The structural metrics change once more from a dense, well interconnected network influence by a key

127

concept (i.e. Major A/B classification roads) to a network which is less clustered (12.75) and larger (diameter = 13). Rather than a focus on a critical SA related variable, the awareness seems to widen.

Country Roads: Driver SA in this environment is dominated by road related features, in particular the geometry of the road such as 'corner' (100%), 'hill' (69%) and 'bend' (60%). Linked to this is a greater awareness of the vehicle's behaviour in this environment, as seen in the emergence of 'gear' (91%) and 'third [gear]' (60%), along with the compensatory behaviours required, e.g. 'slow' (89%). Interestingly, country roads feature the largest number of concepts which lie two standard deviations above the mean, indicating that the content of SA is comprised of numerous highly relevant elements. This is borne out by the lowest score for mean centrality (10.82), which is indicative of a spread of equally important concepts rather than one or two highly important ones. An interesting feature of the structural measures is that this spread of important concepts is not necessarily well interconnected (i.e. density is second lowest at 0.06). This seems to indicate a more discrete set of temporal events, ones that are not necessarily linked to each other in the same way that concepts are inter-related in other situations.

Residential Roads: Unlike country roads, residential roads have the fewest number of concepts two standard deviations above the mean. Unlike major A/B classification roads where a single concept is dominant and interconnected, three critical variables emerge: other cars (100%), behind (50%) and pull [out or away] (38%). The focus seems to have switched from concepts related to the road infrastructure to instead deal with interactions with other vehicles (e.g. avoiding parked cars) and aspects related to vehicle control and manoeuvring. Interestingly, whilst the content is different, the structure of the network is similar to motorways (diameter = 13; density = 0.07; mean centrality = 14.11). This would also seem to suggest that a number of distinct key concepts are active in driver's SA simultaneously.

Junctions: Driver SA in this context is characterized by check[ing] (100%), clear (71%), indicating (69%) and mirrors (48%). Again, whilst the road infrastructure itself is cueing these elements, driver awareness seem more focused on other traffic than on features of the junction itself. That said, throughout the verbal transcripts words and

phrases like 'T-junction', 'roundabout' and 'slip road' occur quite frequently. The key point seems to be that it is the behaviour of other traffic in these situations that is more important for driver SA. The distance across this network is small (9) with density (0.07) and mean centrality (12.01) both moderate. These characteristics are similar to those for major A/B classification roads, again suggesting a focus on specific critical concepts which influence the network as a whole.

4.9.2 Speed Envelope

The on-road study paradigm grants an opportunity to tentatively explore how the content and structure of driver SA relates to actual driving behaviour. In an imaginary, and implausible, case of drivers sticking rigidly to the posted speed limits (and being able to do so regardless of road and traffic conditions) then the time to complete the 14.5 mile course would be exactly 20 minutes. In the event, drivers completed the course in a mean time of 26 minutes and 29 seconds (Min = 23:10, Max = 32:56, SD = 1:50) at a mean speed of 31.08 mph (Min = 26 mph, Max = 35.54, SD = 3.22 mph). A null hypothesis based on a literal interpretation of posted speed limits is too simplistic.

More tenable is that the inevitable differences in posted versus actual speeds should show no marked contingency on road type. Figure 4.18 show the difference between adopted speed of the drivers and the posted speed of the road. This supposition can be upheld based on the control measures taken in the study, and that no drive was significantly impeded by traffic, weather or other external factors. Table 4.3 shows that the biggest absolute difference in posted versus actual speeds occurs on major A/B classification roads, with a mean difference of 14.8 mph representing 32.8% of the posted speed limit for that stretch. This is followed by residential roads, which despite showing a small difference of 6.6 mph actually represents 28.3% of the posted limit. In other words, on average drivers were electing to travel at approximately 23 mph instead of the permissible 30mph. At the opposite end of the speed scale motorways, with a posted limit of 70mph, showed virtually no difference in posted versus actual speeds (0.8 mph / 1.1%).
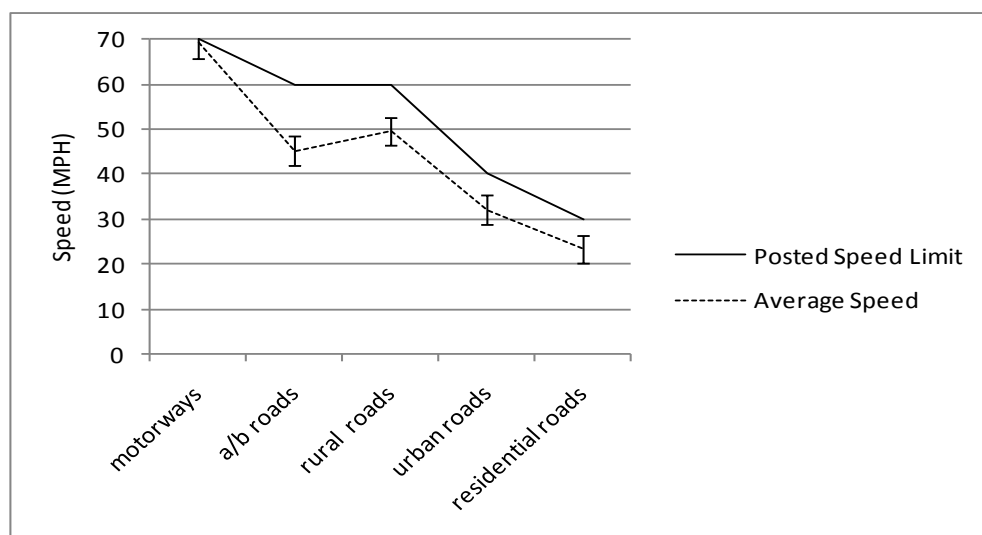
Figure 4.18. Difference is Posted and Adopted Speed

Table 4.3. – Difference between Mean Actual Speeds and Posted Speed Limits

Road type	Motorway	A/B	Rural	Urban	Residential
Difference (posted vs. average)	0.8mph	14.8mph	10.4mph	7.9mph	6.6mph
% of posted speed limit	1.1%	32.8%	21.0%	24.4%	28.3%

These findings are consistent with the discussion above on the self-explaining nature of motorways, and lend a degree of construct validity to the study. Clearly this is an environment in which the context is well matched to driver expectations and intentions, and is an effect that can be readily detected in on-road settings. The data also agrees with the extant work on SER in non-motorway situations. Here the context appears less well-matched to driver expectations and intentions, and adaptations in speed seem indicative of this. As noted above, the marked drop in actual versus posted speeds for major A/B classification roads seems to be associated with SA that is also orientated around the concept of speed. Again, this lends a further degree of construct validity.

4.10 Conclusions

The aim of writing this chapter was twofold. First, it is fundamental to determine the SA needs of a system before proposing to design a SA rich system. It is also essential to determine the SA needs of a system even if the aim is only to enrich the existing system by proposing simplistic human factor guidelines which is the current case. Chapter 4 has defined driver's information needs with a blend of requirement analysis and an on-road study. SA requirement analysis has pinpointed the information needed, defined where it comes from, how it is integrated by the driver's existing schema and how it can be used by designers and engineers.

From SA requirement analysis it was found that 48% of driver SA is extracted from road environment. This is supported by the exploratory study which found 42% of driver SA is assessed from road. Requirement analysis also proved that drivers rely on multimodal feedback for safe and efficient driving. Not only is visual feedback overrated but only a small percentage of information gained through visual stimuli is processed to drive. This is a useful finding for SER which focusses on the visual characteristics of the road to influence drivers (Mackie et al., 2013). Multimodal feedback also increases driver SA (Walker, 2002). Therefore feedback from more than one source is essential for driving.

Chapter 4 in its second part focusses on the relationship between SER and SA. The second half of chapter 4 describes what the driver actually gets from a real life road. Secondary data analysis in Chapter 4 also serves as a pilot study. The results of the pilot study have been encouraging enough to establish the link between SA and SER. In other words it is Chapter 4 which serves as a "valuable adjunct to the developing strand of work in SER" (Walker et al., 2013, p. 27). Chapter 4 fulfils the second objective of this thesis which is to investigate sources of driver information, mode of feedback and evolution of driver information with the increasing complexity of driving.

After the analysis of 1100 driving tasks drivers are found to be extracting highest amount of information from the road. However, if analysed from a micro-level the sources of information vary across various task levels. At a very basic level this provides civil engineers some basic information to work with. For example, a deeper

SA requirement analysis involving the participation of road users will point out the source of SA requirement for other users, and the point of conflict between the two. The requirement analysis here is undifferentiated and looks at the total driving task from a strictly analytical perspective. To deepen our understanding of how real-world environments modify this information requirements study 2 was performed.

The most cognitively demanding tasks (specific and general) rely on the road environment for SA information needs. These types of tasks are dictated by the road environment, such as 'emerging into traffic from other side', 'deal with crossings', 'leave junctions', 'deal with different road hazards', etc. The highest-scoring concepts in motorway conditions are also rooted in the road environment, such as 'lane', 'outside', 'road' and 'traffic'. In support of the primacy of the road environment it can be seen that in the most overtly designed total environment engineers currently provide, the motorway is regarded as the most cognitively compatible and self-explaining in nature. This can be attributed to the design of motorway which matches with driver expectations. The road environment of a motorway is richly designed to supply driver with information he/she needs.

The key insights offered by the analysis of results from naturalistic study include a theoretically robust method of extracting cognitively salient features from road environments. This goes beyond the current state-of-the-art related to picture sorting tasks, questionnaires, and other indirect methods, to instead perform the study with high levels of ecological validity on the road. From the endemic features that have been extracted, it is clear that a strong systemic relationship exists not just between drivers and road environments, but also amongst other drivers and their vehicles. It appears that road design can influence drivers both directly (in terms of how they perceive and behave in response) and indirectly (in terms of how it affects other drivers, and how in turn other drivers behave in that context).

Chapter 5 : Semi-Naturalistic Driving Study Method

5.1 Introduction

Naturalistic study design has been defined for the purpose of this study for several reasons. First is drivers drive on a real life road without any probe during their drive. This offers complete ecological validity and provides an insight to driver-road interaction. Secondly, this property of interrelationship is of particular interest to this thesis as SER and DSA are based on studying the road environment as a system of continuous interaction between driver, vehicle and road environment. Furthermore, Chapter 4 proved experimentally that road environment is the most important source of driver SA. Therefore, it is pivotal to choose a method which can elicit this relationship. The on-road study is the only method available to do that. Thirdly, it offers experimenter a wide range of choice of performance and process measures to collect data on driver behaviour. Another advantage of using NDS is it provides exposure to risk related behaviour (Schagen et al., 2012). For example, abrupt breaking, less headway, tailgating, etc. this behaviour would mean poor interaction between environment and driver and an indication of potential danger zones on a road.

Naturalistic driving study (NDS) provides insights into everyday driving behaviour (Schagen & Sagberg, 212). NDS is the state of the art in traffic safety research (Schagen et al., 2011). NDS is of two types: on-road study and on-site study. Typically in NDS (on-road) drivers drive their own vehicle on a predefined route with advanced technology recording driver behaviour unobtrusively (SWOV, 2012). In an on-site based observation, observation equipment is used alongside the road rather than inside the vehicle (SWOV, 2012). Road user behaviour can be studied at specific locations through on-site observation. On-road version has gained much prominence and been used by 100 car naturalistic study whereas the on-site version is not in use much.

Carsten et al., (2013) reported a different categorization of on-road studies. According to Carsten et al., (2013) on-road studies can be further divided into controlled on-road, field operative tests (FOT) and naturalistic studies (NS). Controlled on-road studies are the smallest group and they aim to study driving behaviour and performance as affected

by design, distraction, experience, fatigue, etc. (Carsten et al., 2013). These studies are very short and last for a few minutes to a couple of hours. FOTs have been used for over twenty years. They are mainly evaluative in nature, testing the impact of driver assistance systems on driver behaviour and safety. FOTs last from a couple of hours till a few days. Finally NS are diagnostic in nature and collect data without any control for days and months (Carsten et al., 2013). NS aims to investigate the relationship between different variables leading to accident causation; such as mobile phone use and accident causation (e.g., Dingus et al., 2006). Some very popular NS are the 100 car naturalistic study where data was collected for 1 year by Dingus et al., 2006) and the 60 vehicle 2 years long study conducted by Uchida et al., (2010). A new hybrid method, FOT and NS, is also on a rising trend (e.g. Hickman & Hanowski, 2010). In light of this information the study described in this chapter can be categorised as controlled on-road study or as a semi-naturalistic study. Semi-naturalistic implies that the study is conducted to explore a diagnostic relationship between road design and SA and furthermore speculate on the relationship between SER and SA. However, the study is highly controlled and drivers drive on a pre-defined route with the experimenter. Therefore, on the basis of these practical aspects this study is best termed a controlled-on-road study.

Chapter 4 has put forward an argument that SA is a valuable adjunct of SER. Through its exploratory setting it is able to apply the sophisticated method of DSA into the real world setting and mine endemic features from the road environment. This is consistent with the aim of this thesis. It is now fair to scale up this exploratory study to achieve the rest of the objectives set out in Chapter 1. This chapter is written to overcome the limitations of the study of the second half of Chapter 4 and to proceed with the large scale naturalistic study. Therefore Chapter 5 details the setting up of the on-road study, sample selection, route selection and materials used. Finally, it ends with the precautions to be followed in order to complete the study successfully.

5.2 Methodology

5.2.1 Research Design

Road type will serve to be the independent variable (I.V). I.V has two levels; M8 and A roads. M8 being the motorway will act as a control condition as it is found to be richest in SA and closest to being a SER in the pilot study in Chapter 4. A71 acted as the experimental condition. Driver behaviour acted as a dependent variable (D.V) as the driver negotiated the road. D.V was obtained through verbal protocol and situation awareness rating technique (SART). Workload (NASA TLX), driving style (DSQ), traffic flow and weather conditions were controlled. Before starting the experiment as part of administration participants were sent a DSQ to detect their driving styles. To control weather and traffic all runs were conducted between 1330 to 1530 hours. Furthermore, all the participants were played back an existing police driver audio recording to give them an idea what is expected from them during the drive in the verbal protocol. Since the same recording was played to train 34 drivers it acted as an experimental control itself. This helped to control any confounding variables.

Following completion of a drive the verbal protocol data was subjected to an automated textual analysis. This resulted in knowledge objects and their structure which characterise driver situational awareness. This is discussed in Chapter 7. Following this a network analysis was conducted. To determine popularity and connectivity of concepts a social network analysis was also performed. These are explained in Chapter 7 and have been performed in the pilot study described in Chapter 4. An A-B-B-A experimental design which compares motorways (A) and A roads (B) in a counterbalance sequence is followed in order to remove order effects.

5.2.2 Participants

35 participants voluntary participated in this study. This number was selected so as it mirrors the everyday driving population of Scotland (Figure 5.1). This implies 20.461 ~ 20 males (to make up 58.46% of 1096) and 14.528 ~ 15 females (to make up 41.51% of 1096).
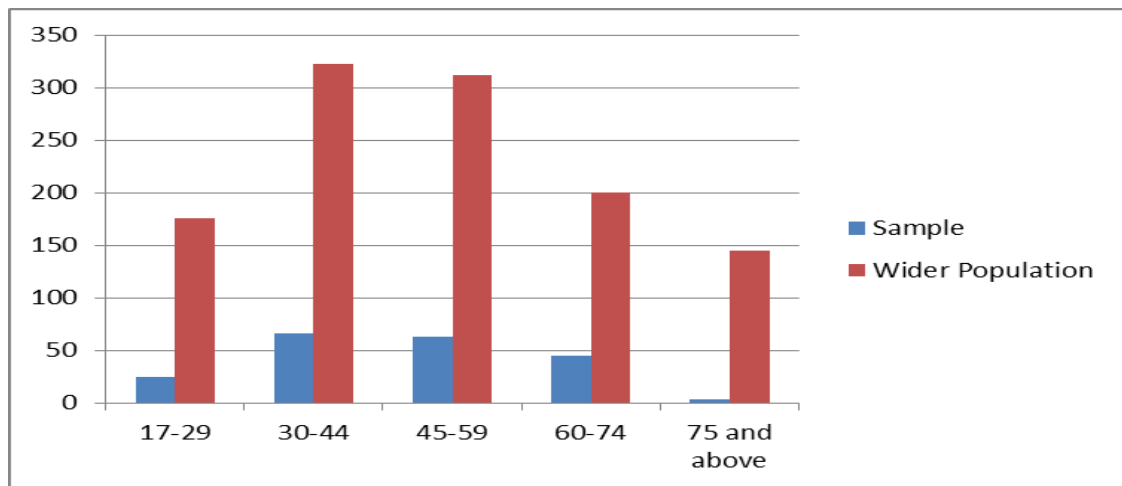
Figure 5.1 Similarities in sample with the wider population

Table 5.1 Details of the sample selected

Age Group	Males	Females
17-29	3	2
30-44	6	5
45-59	6	5
60-74	4	2
75 and above	2	1

For the purpose of these study participants with full driver licence, valid MOT and comprehensive insurance with business use were approached. Drivers were recruited by various means such as advertisements, word of mouth and emails. To encourage participation participants were compensated in the form of high street vouchers for their time and fuel. To comply with health and safety regulations outlined for this study registered drivers of Heriot-Watt were asked to participate first. Registered drivers include people of all age groups ranging from staff to students. These registered drivers were first contacted by Heriot-Watt risk officer through an initial email. In case of non-registered drivers, participants scanned a copy of their licence, insurance cover and MOT certificate to the experimenter. Once they were checked and approved by the risk officer, a date and time was scheduled with the participant for the drive.

5.3 Materials Required

5.3.1 Vehicles which people will drive

The vehicles used in this study were required to have complete comprehensive insurance cover with business use. The vehicles were checked personally by the experimenter for, a valid tax disc, a valid MOT certificate and smoke-free environment. All these three are requirements of DVLA for a vehicle to drive on road. Owners of the vehicle were also required to hold a valid driving licence, should not have had any major accident in the past. Drivers who were detected as "rapid boy racer" by DSQ were excluded and not contacted further for participation. As a backup plan in case of not getting enough vehicles with business insurance cover, there was a provision for driving a hired vehicle with an extended test drive which was insured by the Heriot-Watt Estate Office. This option was not used as the required number of participants had their insurance covered.

5.3.2 Test Route

The route was a 22.5 miles long and based in the West Edinburgh Area. The training route started from Boundary road east and goes till Research Avenue North opening to Hermiston Part and Ride. This was a 1.3 miles long Heriot-Watt University campus enclosed road which was used for practice and 'warm-up'. The motorway section was the M8 from junction 1 to junction 3, junction 2 being the exit point to M9. This was a 10.6 miles long road. It is built to UK motorway specifications with two 3.65 m wide running lanes and a hard shoulder etc. The posted speed limit for the entire length was 7 mph. A71 had 2 sections; the rural A road section was the A71 between Hermiston Park and Ride to Lizzie Brice's roundabout (9 miles). From the roundabout there was a 2.9 miles long urban dual carriage A899 road till Livingston roundabout. So the A road drive was 11.9 miles in total.
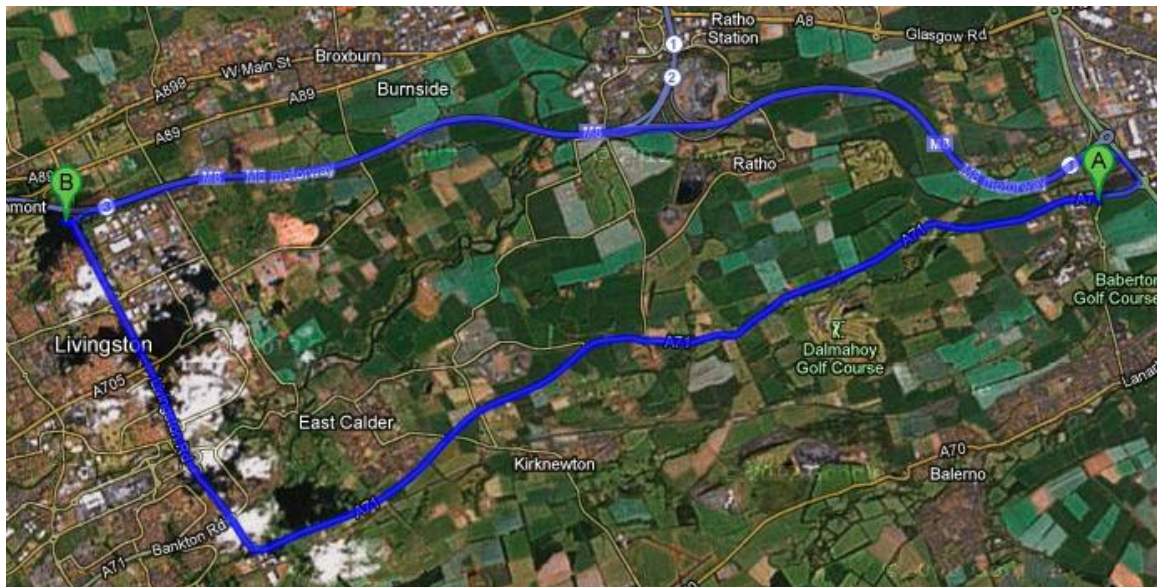
Figure 5.2: Experimental track

5.3.3 Voice recording device

Sony ICD-PX312 digital voice recorder was used for recording verbal protocols of the participants. It stores files, and can play back files in a compact mp3 format. The 2 GB internal memory allows storing 856 minutes of dictation. It can be synchronized to be used with any computer with the help of an usb cable and with any media player.

Figure 5.3: Sony Voice Recorder

The recording was transcribed in excel systematically and separately for M8 and A71. For the purpose of further analysis and intermediate step between recording and analysis was writing the verbal protocol obtained in Microsoft excel.

5.3.4 Leximancer™

It is an automated text analysis tool which was used to generate three forms of output. First, it produced a concept map, second it yielded different concepts in the transcript upload and thirdly, it gave relevance of each concept in terms of percentage. In this case, concepts were extracted from the semantic units (which are sentences or dialogues of drivers) in the text file; and then semantic networks were created. Leximancer™ does this automatically once the file is uploaded. Leximancer™ ™ has been used extensively in previous studies. It has provided insight into organizational change (Rooney et al., 2010), intergroup communication (Hewett et al., 2009), analysis of web content (Coombs, 2010) and large scale meta-analyses of themes within scientific journals (Cretchley et al., 2010). It is assumed that the use of this tool in analysing verbal protocol will provide interesting relationship between few concepts in real-life transportation domain.
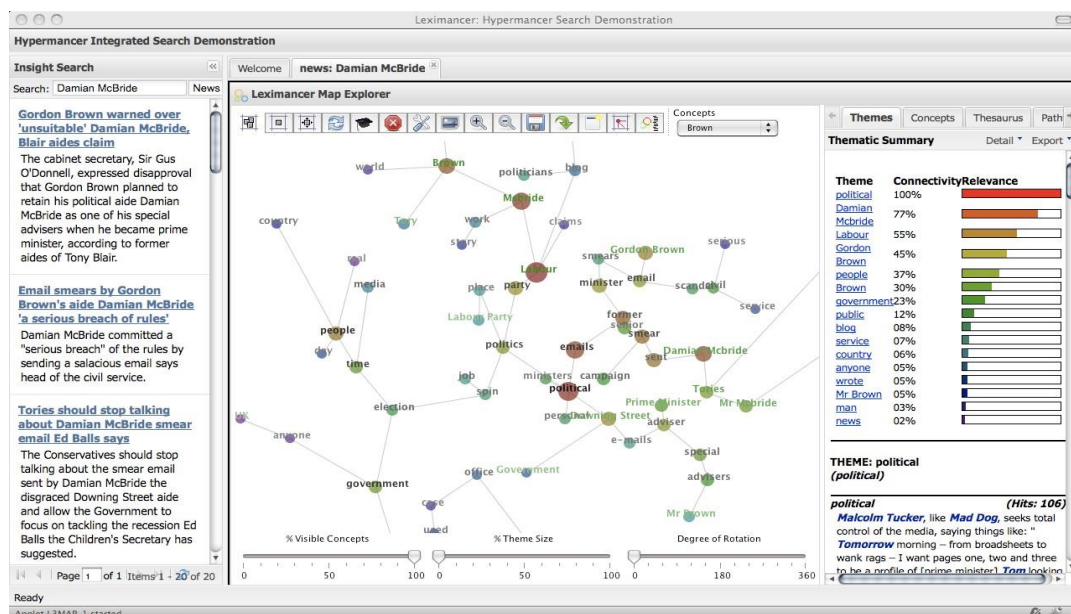
Figure 5.4 : Sample Leximancer™ Output

5.3.4 Agna

It is a free social network analysis tool. Density, diameter and centrality status can be determined using Agna. Diameter will tell how well the two concepts are related and centrality status will reveal the popularity of a concept.
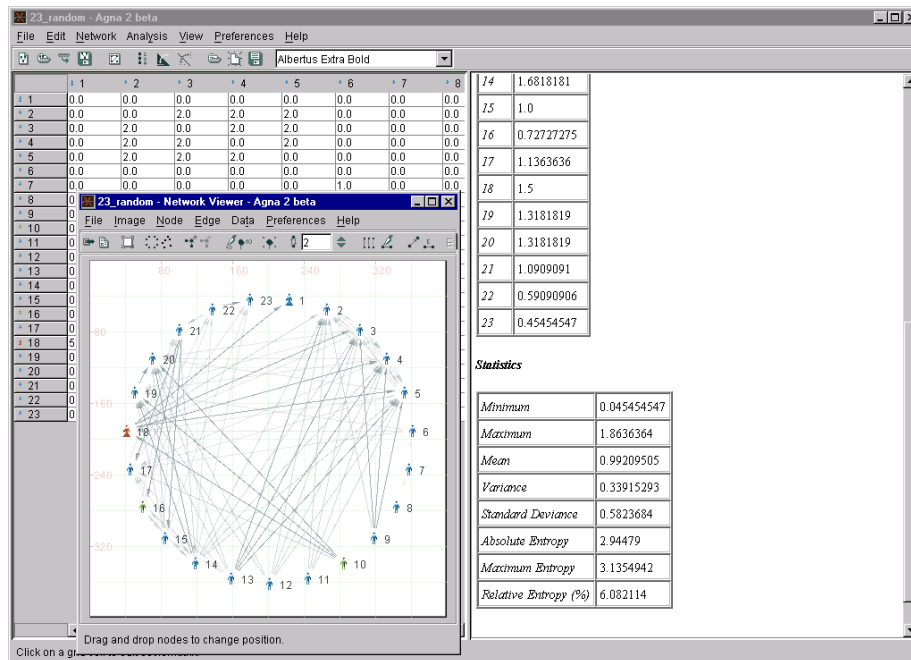
Figure 5.5: Sample AGNA Output

5.3.5 Situation Awareness Rating Technique, Taylor, 1990 (SART) (Appendix A)

This is an easy to administer self-rating technique, typically used with SAGAT. In conjunction with SAGAT, SART is the most popular tool for assessing SA. It measures 10 dimensions of situation awareness namely;

Familiarity of Situation	Focussing of attention
Information quantity	Instability of Situation
Concentration of attention	Complexity of the situation
Variability of the situation	Arousal
Information quality	Spare Capacity

140

Apart from the 10 dimensions version a quicker version of SART is also available. It is called the 3D version. The 10 dimensions are grouped into 3 categories. They are:

1. Demand of Attentional Resources: Mix of complexity, variability and dynamicity of the situation (Stanton et al., 2005)
2. Supply of Attentional Resources: Adequate mix of attention, concentration, mental capacity to meet the demands.
3. Understanding of the situation: Mix of quantity and quality of information in addition to familiarisation of situation (Stanton et al., 2005).

After the participant has finished rating, the participant SA is calculated using the following formula:

$$S.A = U- (D-S)$$

U: Summed Understanding

D: Summed Demand

S: Summed supply

It does not require any extended customization (Endsley, 1998) and can be readily applied to the desired domain. The SART ratings are often correlated with subjective measures of workload and performance of pilots to determine the efficacy of cockpit designs (Endsley, 1998). Out of all the measures of SA, SART perhaps has the highest number of validation studies associated with it.

5.3.6 Driving Style Questionnaire (DSQ) (Appendix B)

Given by French, et al., 1993 and developed from decision making questionnaire (DSQ). It has 15 items and is in the form of a rating scale. The participants can select their response from 1 (being never) to 6 (very frequently) for each item. It classified the driver as having one of the following driving styles; calmness, speed, social resistance, focus, planning and deviance. Scoring of DSQ is very simple. There are some questions allocated to every style. For example, Q1, 2 and 3 are for speed, 14 and 15 are for

deviance and so on. Once completed by the participant each category can be scored separately.

5.4 Procedure

Participation was absolutely voluntary and participants had the right to withdraw at any point during the study. They were compensated for their time and effort in the form £20 cash prize. Formal recruiting of participants only began after obtaining the ethical approval of the Heriot-Watt University's ethics committee.

Participants who expressed interest after seeing the advertisement were sent out a driver declaration form. The university's driver declaration form (Appendix C) were filled and returned with a copy of MOT certificate, driving licence, and insurance certificate to the experimenter. These documents were further subjected to be checked by the university's risk officer. Once all these checks are cleared all interested participants were provided with an informed consent (Appendix D) and asked to fill DSQ two days prior to the date of the run. There was also an alternative of hiring a Vauxhall Agila for participants who are unable to provide the requisite documents. The hired vehicle if used will be insured under the guidance of risk officer.

The experimenter accompanied all the participants in case they needed additional information about the route or faced any problem during their drive. A secondary aim for accompanying the participants was to collect any additional useful data through informal interviews and observation. The presence of experimenter in the car did not have any effect on driver behaviour. Since most of the participants were employees of Heriot-Watt and lived in that area they were very familiar with the route. Thus, the experimenter was completely ignored by the participants. Moreover, none of the participants hesitated in producing verbal protocol indicating presence of experimenter did not pose any workload on the participant.

The experiment run started following the completion of trial run in Heriot-Watt University campus (Figure 5.6). Participants started driving on A71 and came to a halt at BP Deer park at the end of A899 (Figure 5.7). During their route on A71 drivers negotiated 2 major roundabouts; Lizzie Brice and Oakbank, switched speeds from 40

mph on A71 at the beginning , dropped to 30 mph through their drive in Wilkieston village and then drove at 50mph until they came to a halt at the deer park. During the drive drivers also passed through numerous intersections and experienced roadworks once. At the halt drivers were asked to fill up NASA task load index and SART. The motorway (M8) phase of driving started from BP Deer Park (Figure 5.8). On M8 drive drivers negotiated Livingston East roundabout at the beginning and drove through junction 3 and 1. On M8 drivers drove at a steady speed of 70 mph. At the end of M8 drivers were required to go through the 6 lane city Byepass. The last phase of the driving was from City Byepass until Heriot-Watt University Car Park. This phase had Calder road roundabout and speed variation from 40 mph to 50 mph. On reaching Heriot-Watt drivers were asked to fill SART and NASA task load index again. After collecting all the measures the experiment was terminated.

The following instructions were given to the participants:

"Thank you very much for coming to Heriot-Watt for being a part of my thesis. My thesis is on how and why drivers differently on A roads and motorways. So, to get to know that I need you to drive you on a route which have been agreed on by the university ethics committee. Basically, you need to speak while you drive and I will be recording what you are speaking. It is something like this [sample recording]. In other words, I need to know what you are thinking so 'think aloud'. Everyone thinks of the next manoeuvre he/she will be doing but they don't talk to themselves loudly about that. So drive as you would drive, ignore my presence beside you, keep talking about the road and traffic, have a safety first attitude and start the drive. The important thing is to keep talking. Having said that, I understand this is not something which you do every day so we can first go for trial run inside the university campus and then go on to the main road. The route of the experimental drive is this [show map]. So you need to start from the end of research park, onto A71 and then all the way to Livingstone roundabout. At the roundabout, (it has a big A like structure) pull over to BP deer service station. During the stop you will be asked to fill two questionnaires. After that you start driving on M8 junction3 which is the first roundabout form your halt. The M8 drive ends on reaching the Edinburgh City Byepass. This actually ends your motorway drive. Then you drive back to Heriot-Watt, and two more small questionnaires will be administered on you. Once your responses are received by me, you will be free to go

with your vouchers. At any point if you will uncomfortable or uneasy we will stop the experiment. I will direct you if you need directions"

5.4.1 Route Traversed by the Participants

Figure 5.6: Trial track.

Figure 5.7: A71 from Hermiston Park (A) and ride to Lizzie Brice roundabout. A899 from L-B roundabout to Livingstone roundabout. Pull over point BP deer park service station (B).

Figure 5.8: M8 from Livingstone roundabout (A) to Hermiston Park and Ride (B) (10.6 miles). M8 section precisely starts from J3 near Livingstone roundabout and ends at the slip road, right after J3 (9.8 miles).

5.5 Data Analysis

The data collected was then subjected to a detailed analysis. First, 2 SARTs, 2 NASA TLX and DSQ were scored manually. Secondly, transcription of verbal concurrent protocol was conducted in an excel sheet. This yielded sentences similar to what we had in SA requirement analysis. So, this transcribed data was further coded into visual, auditory, tactile for feedback and car, other traffic and environment for sources of information. So responses under all these categories were counted and scored separately. To ensure reliability 2 independent raters will rate 10% of the transcripts the inter-rater reliability will be calculated using SPSS. Once, satisfactory inter-rater reliability was achieved the resultant was used to derive formative SA. They were compared with the results of requirement analysis and the pilot study done in the preceding chapter.

Third step of analysis was electronically done with Leximancer™. The transcripts of motorway for each driver was copied and pasted in separate txt files. Similarly, transcripts of A road was copied from excel and pasted in txt files. This means 70 txt (35 M8 and 35 A71) files were obtained. Afterwards, in the Leximancer™ portal these files were uploaded. So, Leximancer™ created semantic network for each file uploaded. In addition to semantic networks, it also provided a list of concepts with frequency and themes. All the concepts obtained were inserted in AGNA. This reveals how strongly

145

two concepts related and which ones are more important than others. The end result of the entire analysis was some concepts for every participant unique to each category of road they drove on.

5.6 Precautions

1. All runs were done in clear dry conditions. Mild rain and drizzle are tolerable.
2. A map of the route was kept handy. In case participants were not familiar with the route a map was shown to them prior to the drive.
3. Every driver was checked for MOT certificate, insurance certificate and driver's licence.
4. Before leaving the university campus a text message was sent to a member of Human factor (HF) Research Group at School of Built Environment (SBE).
5. Experimenter was supposed to stay silent during the run to avoid distractions and increasing mental work load.
6. Consent form was sent 2 days before the actual run so that participants can read and re-read the form carefully and understand each sentence before giving their consent.
7. The trial run ensured if drivers are fit to go for the experimental run as was the driving style questionnaire
8. A text message was sent to the HF research team upon safe completion of the run.
9. Emergency contact numbers (for taxi transport, emergency) was be carried on the drive in case of unexpectedly having to stop or abandon.

5.7 Summary

Chapter 5 has presented the semi naturalistic driving study set up for this thesis. The purpose of this chapter is to overcome the limitations of Chapter 4 and use the results to build on DSA for describing SA in road transportation context. Chapter 4 achieved what it was supposed to achieve as a pilot study. It established that road is the primary factor in driver SA. It also proved that SA is a useful extension of SER. Despite this pilot study in chapter 4 had some limitations. First was because it was pilot study it had a rather small sample size (n=12). It used a homogeneous sample of male drivers only. Highest number of everyday drivers in Scotland belongs to the age of 41-49 (High level

Summary of travel and Transport, 2011). No drivers were above 44 years in the pilot study. These three points reduced and restricted the representativeness of the sample. Thus, this study is designed to overcome these limitations through a representative sample of 35 participants. All the participants used their own vehicles and drove without any technical problems on the route. One verbal protocol was discarded because of its poor quality. Therefore 34 verbal protocols were analysed for structure and content of SA on A71 and M8. The implications of these results are discussed in Chapter 9 and Chapter 10.

A recurring theme is this thesis is the systemic view of the road which focuses on the above mentioned relationship. DSA selected in chapter 2 asserts that road is a complex socio-technical system where drivers, vehicles and road are a part of. SER also states that the road can generate key messages. In order to determine a relationship between SER and SA it is essential to determine how roads and drivers are connected. In other words, how these two interact. That is, to explore the communication links between the different agents. NDS is the only research method which enables to do so through the rich data it collects.

Driving behaviour is traditionally researched using driving simulators, self-report measures, instrumented vehicles, analysis of crash statistics and in-depth analysis of accident data (Eonink et al., 2014). In these methods like any other psychology experiments there is a trade-off between experimental control of outside variables and ecological validity (Brewer & Crano, 2000). Ecological validity is the extent to which the materials, methods and results match with the real world that is being examined (Brewer & Crano, 2000). SNDS overcomes this limitation through its design. Drivers do not always do what they say they do (Meyers et al., 2011). Moreover, not all simulator study results are translated to real world traffic (Olson et al., 2009 cited in Regan et al., 2012). Hence, SNDS data helps to validate the self -report measures and simulation studies (Regan et al., 2012). It offers 100% ecological validity and studies driver behaviour in its basic state. SNDS data is an excellent means of evaluating the effectiveness of new and existing countermeasures such as new road design treatments, advanced driver assistance systems, training programs, etc. to improve driver safety (Regan et al., 2012). Using SNDS will enable to further investigate SA and SER in A71 and M8.

The 34 transcripts obtained from SNDS were analysed for DSQ, NASA TLX and SART results with Leximancer™ and Agna. They are discussed in Chapters 6 and 7.

Chapter 6 : Direct and Indirect Measures of Driver Behaviour

6.1   Introduction

The aim of this thesis is to lay the foundation for user centred roads from a distributed situation awareness (DSA) perspective. That is to say, purposefully design roads to interact with motorists and other road users. From the preceding, Chapter 1 identified a research problem which this thesis will operate within, Chapter 2 through a careful review selected distributed situation awareness model to explain situation awareness (SA) for the current problem, while Chapter 3 proposed a methodology to tackle the problem identified.   The proposed methodology was tested using secondary dataset in Chapter 4. The results were encouraging to proceed with a full scale naturalistic driving study (NDS). Therefore, in this respect Chapter 5 detailed the process and setup of naturalistic study undertaken. Chapter 6 will discuss the results from situation awareness rating technique (SART), NASA task load index (NASA TLX) and driving style questionnaire (DSQ) obtained from NDS. Chapter 6 and 7 together justify the fourth objective of this PhD which is, extract cognitively salient 'endemic features' required by SER approaches using DSA methods and validate them.

The results of NDS will be discussed in three broad sections. First one being the control measures i.e. things which were measured that gave us confidence that what we were observing were really artefacts of our manipulation of road type rather than peculiarities of the sample. In this study the control measures were driving style questionnaire (DSQ by French et al., 1993), NASA task load index (NASA TLX) and speed adopted by the drivers. The second section of results will discuss the results of self-reported situation awareness. In this experiment, only one such measure was used which is situation awareness rating technique (SART by Taylor, 1990). Third and the final section will derive meaningful conclusions from the verbal protocols in chapter 7.

The procedure followed to obtain these results is explained in full in Chapter 5. 35 drivers voluntarily participated and agreed to drive a 22.5 mile route. The experimental route starts from Hermiston Park and Ride. The A road segment is 11.9 miles long route till BP deer park. The drive has two halts. One at the end of A71 at BP Petrol station

and the other one at the end of M8 junction 3 in Heriot-Watt University. In both the halts, two assessments i.e. NASA TLX and SART were administered. During the driving phase drivers provided a concurrent verbal protocol.

The results of pilot study in Chapter 4 have informed the design of this experiment. Lessons learnt from pilot study were motorways were most cognitive compatible and incompatibility increases as roads become minor (Walker et al; 2013). Thus M8 served as a control condition. Some of the features of SA rich roads as identified by Walker et al., (2013) where presenting data in a manner which makes understanding and prediction easy for road users, provide critical cues to capture attention during critical events and many more. The exploratory study strengthened the researcher's confidence to proceed into full scale testing. Thus, a larger sample of 35 registered drivers with business insurance on their vehicles was selected.

6.2 Time
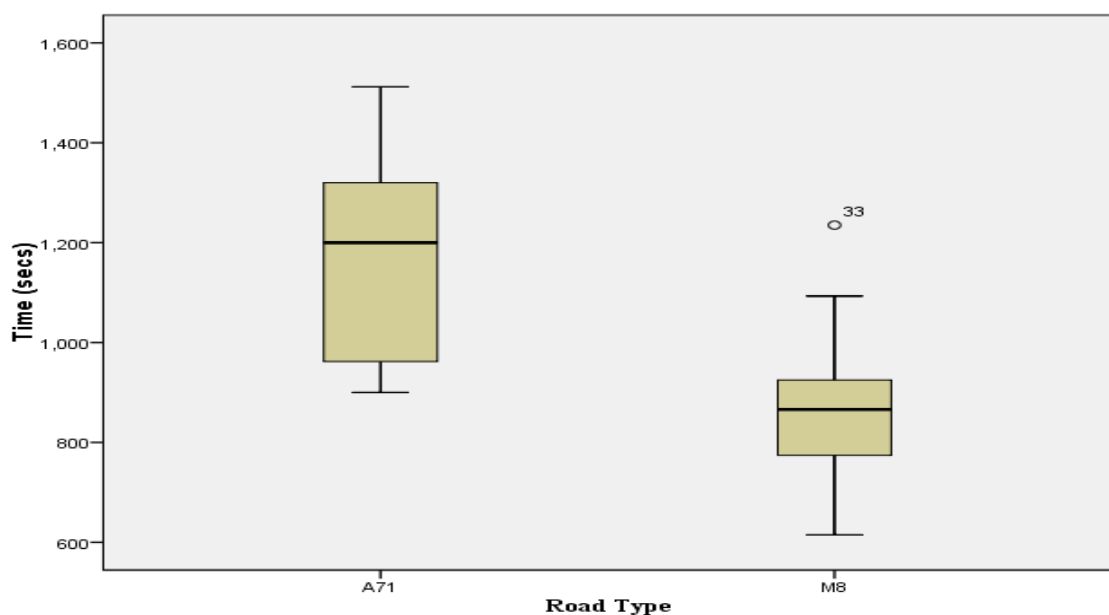
Figure 6.1 Time Taken by Drivers on A71 and M8

Figure 6.1 shows that participants on average took longer to drive on A71 than M8. Drivers drove at an average speed of 38.062 mph on 11.9 mile long A71 and took 19.358 min. On the other hand drivers drove 10.6 miles long M8 at an average speed of 63.393 mph in 14.460 min. Therefore, drivers took approximately 5 min (4.898 min)

less on M8. The Figure also indicates that A71 has a higher variance in time taken than M8 (A71 S.D = 3.18 min, M8 S.D= 2.09 min). Hence, drivers have completed their drive on M8 at a similar time and very likely at a similar speed with each other.

6.3 Speed
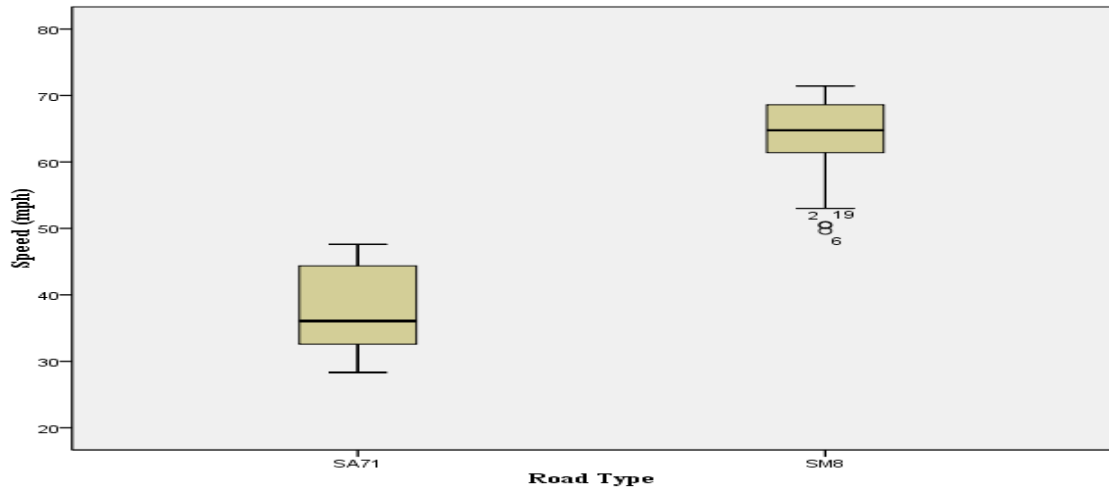
Figure 6.2 Box plot of speed adopted in A71 and M8

Clearly the adopted speed of M8 is more than A71 due to difference in functionality in A71 and M8. Figure 6.2 show that most participants on A71 chose to drive over the average speed of 38.06 mph with S.D of 6.13 mph. The range of speed in A71 is slightly lower than M8 (-2.54 mph). Interesting difference can be seen in variance of speed. The difference in variance between A71 and M8 is >1 mph. However, the difference in variance in time is 346.127 i.e. 5.5 min. This indicates that participants took longer to interpret road sections of A71 while adhering to the speed limit. Since A71 is demanding drivers spent more time in looking of cues and resources which helped them to drive forwards. This is better explained in chapter 7 through network analysis. On the other hand in M8 almost all the drivers adopted the mean speed of 63.39 mph and readily negotiated the route at a same time. The perceived image of M8 is closer to the objective image of the road indicating towards its high SER nature and user centeredness. This also echoes with the results of pilot study in chapter 4.
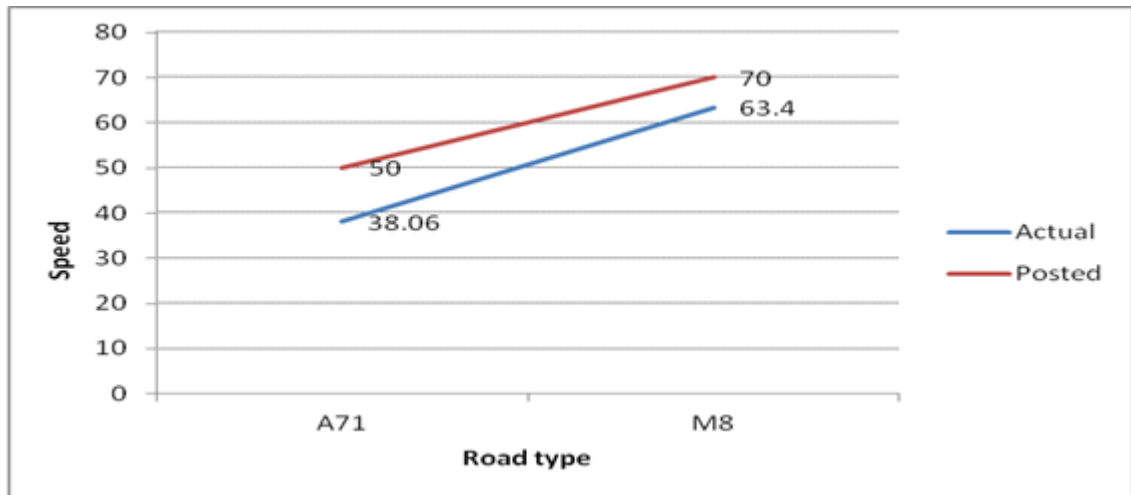
Figure 6.3 Speed profile showing posted speed and actual speed (n=34)

In an implausible condition of drivers sticking rigidly to speed limit drivers should complete the 22.5 miles long road in 23 minutes 30 seconds. In this event drivers completed the route in the mean time of 34 minutes 28 seconds (min = 8 mins 25 seconds on M8, max= 25mins 2 seconds on A road, S.D A71 =3.18 min, S.D M8= 2.09 min). As already proved in pilot study the difference between actual and posted speed limit is higher in A road (11.94 mph) than Motorway (6.6 mph). This suggests the high ambiguity, variability and heterogeneous nature of A71 as opposed to M8. Motorway design supports derived speed and time profiles.

A road which is 11.9 miles long (from Hermiston Park and Ride to Livingston East Roundabout) has poor SER features. Low functionality on A71 can be defined by the worn out road markings, poor road texture and obscured road signs. These are encountered in the initial phase of A71. Low predictability of A road is best shown in A899 (3 miles long) which is the last phase of A road drive. The speed limit of A899 is 50 mph as opposed to the standard speed of 70 mph on dual carriageways. A899 being dual carriageway confused the drivers due to poor standardization and being heterogeneous within a road category. This heterogeneity of A road is further seen in the variable speed limits; 30 mph from Wilkieston to Kilmarnock junction, 40 mph from Hermiston Park and Ride to Ratho Crossing.

The difference in the actual and posted speed limits indicates a cognitive incompatibility between road design and car drivers. The concept of genotype and phenotype schemata is particularly relevant here to understand what cognitive compatibility is. Genotype schema is the likelihood of an individual to behave in a certain way in a situation containing some specific features (Walker et al., 2012). Phenotype schemata are the activated schemata (Stanton et al., 2009). The higher discrepancy in speed limits of A71 means that the existing design of A71 is triggering the inappropriate schemata. This is one of the three schema related problems given by Neisser, 1976.

Although all possible efforts were made to control weather and traffic flow, but, 6 of the participants faced rain and winds in the middle of their drive. Traffic was constant for all participants. Therefore, traffic flow and weather are uncontrollable variables. These 6 participants had an average speed of 55 mph because they had to drive in rains and winds. Thus their speed was hindered by waters sprays from other vehicles, poor visibility and presence of more Lorries.

6.4 Control Measures

6.4.1 Driving Style Questionnaire

Driving style is defined as, "the way individuals choose to drive or a driving habit that have become established over a period of years. It includes choice of driving speed, threshold for overtaking, headway and propensity to commit traffic violations" (Elander et a., 1993, p.279). DSQ as presented by French, West and Elander in 1992 is made of 15 items which ask about attitudes to/engagement in different types of driving behaviour etc. Each item provides with options ranging from never (0) to always (1). DSQ provides insight to 6 driving styles. The subscales are derived through principal component analysis of DSQ with 39.4% variance among its subscales (West et al., 1992). They are speed (Do you exceed the speed limit in built up areas?), calmness (Do you remain calm when things happen very quickly and there is little time to think?), planning (How often do you set out on an unfamiliar journey without first looking at a map), focus (Do you drive cautiously?), social resistance (Do you dislike people giving

advice about your driving?) and deviance (Do you ever drive through a traffic light after it has turned red?). Each of these driving styles has been found to relate significantly with accident rates and behaviour (West et al., 1992, French et al., 1993).

DSQ scores are also found to have correlated with age and sex (West et al., 1992). For example, speed showed a steady decrease (-.35) with age in males and females (6.3) (West et al., 1992). Females although scored lower than males (West et al., 1992). Similarly, social resistance was found to be insensitive to sex (-0.4) and decrease with age (-0.12) (West et al., 1992). Deviant behaviour had negative correlation with age (-0.18) irrespective of gender (0.9) (West et al., 1992). Further to high correlation among the subscales of the questionnaire they have also high test-retest reliability (speed = 0.70, calmness = .53, planning = 0.43, focus = 0.29, social resistance = 0.53, deviance= 0.44) (West et al., 1992).

The DSQ was administered for two reasons; the first is health and safety and the second is to ensure representativeness. For health and safety reasons DSQ was administered to detect if any participant possessed reckless driving tendencies. The ones with such tendencies were excluded from the study. DSQ was given to 35 experienced drivers. The average scores are the highest in focus followed by calmness (Figure 6.4). The results of DSQ on a larger population on 711 drivers (West et al., 1992) show a similar trend (Figure 6.4) to the one derived from 35 participants. Apart from the couple of differences the sample looks very well aligned to the larger population. This enhances the representativeness of the sample and increases the generalizability of results obtained.
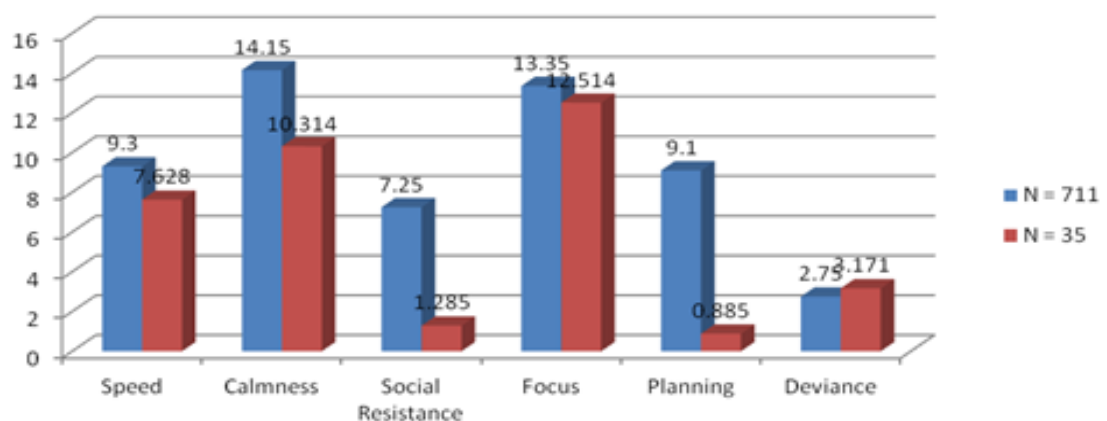
Figure 6.4 Comparison of DSQ scores in N=35 and 711

The large scale (n=711) administration of DSQ has higher number of older respondents. This can account for high scores in planning rather than deviance in Figure 6.4. As mentioned above deviant behaviour is negatively correlated with age. In the sample of 35 drivers only 1 subject was above 60 and rest 34 were between 23 to 56 years of age. This explains the higher score on deviant behaviour rather than planning in Figure 1. High correlation (r=0.748, n= 6, p= 0.87) between two sample sets were found. It means the rise and fall of DSQ results in the smaller sample as well as the larger samples are influenced by same stimuli.

6.4.2 NASA Task Load Index

Workload is the demand on the brain and sensory system such as eye, ear and skin due to the task (Zhang & Luximon, 2005).). Underload leads to reduced alertness and lowered attention (DeWaard & Brookhuis, 1997). Overload leads to distraction, driver inattention, insufficient time and capacity for informational processing and diverted attention from the task (DeWaard & Brookhuis, 1997). High scores on workload is found to be linked to driver impairment (DeWaard, 1996)

Originally given by Hart & Staveland, 1988; NASA TLX is a multidimensional subjective workload assessment tool. It was designed to assess workload in pilots but now they are used for all cognitive scenarios (Less, 2007). NASA TLX was developed through 16 experiments over a period of 3 years (Hart & Staveland, 1988). 13 experiments on 6 male subjects were performed for its validation (Hart & Staveland, 1988). NASA TLX has high test-retest reliability (0.83) and extensively validated. NASA TLX has 6 subscales namely, mental workload (perceptual activity such as thinking, deciding, searching), physical workload (turning, pushing, pulling), temporal workload (refers to time pressure on the task), frustration (irritation, intolerance, stress), performance (related to personal goal accomplishments) and effort (energy spent and effort put in to complete the task). Each subscale is rated on 20 point bipolar scale with a 5 point interval. The overall workload score is the sum of weighed score of each dimension. There are two versions of this scale available paper and online. For the purpose of this thesis paper-pencil version of the scale was used. The first three scales (mental, physical and temporal) represent the demands imposed on the subject. The last three scales (frustration, effort and performance) represent the demands imposed

through interaction of a subject with the task. NASA TLX is very quick and easy to administer. NASA TLX is a sensitive indicator or overall workload across a variety of tasks (Hart & Staveland, 1988). Lesser the score more relaxed and stress free is the drive. NASA TLX was used as control measure in this experiment. It is to check if concurrent verbal protocol has imposed any additional workload on driver making the driving exceptionally demanding.
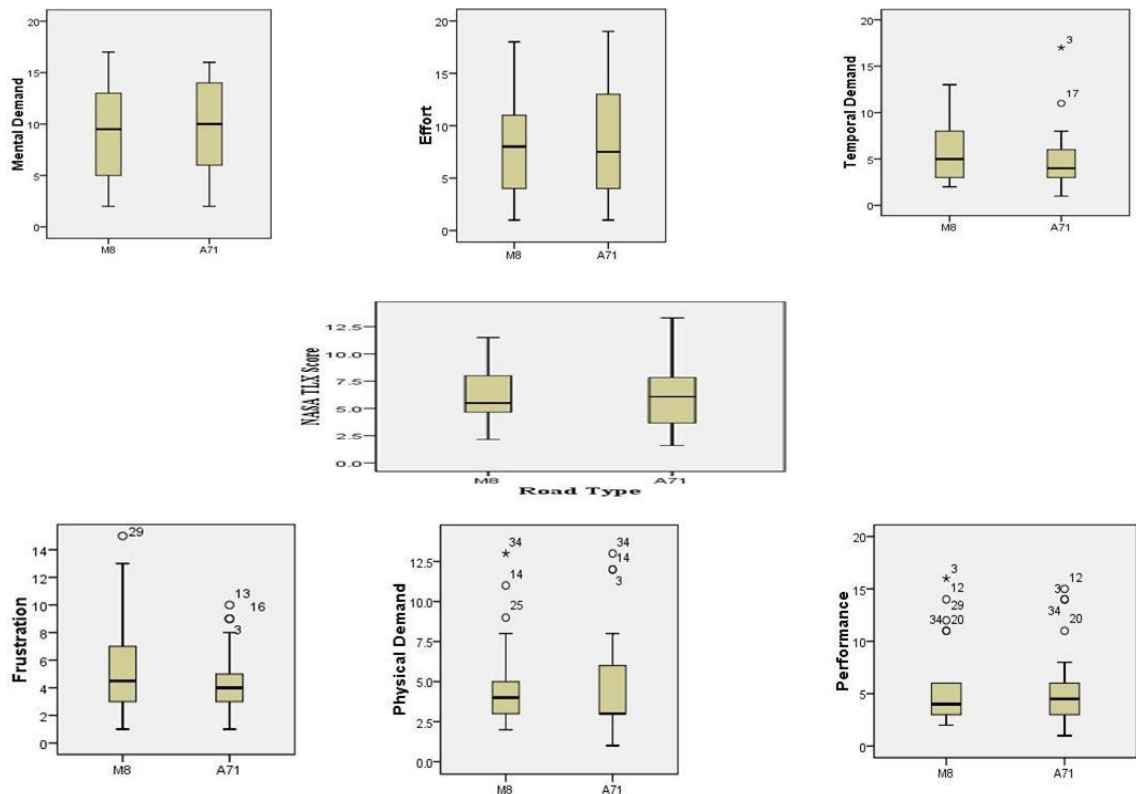
Figure 6.5 Workload Scores across M8 and A71

As seen from Figure 6.5 the workload of participants are neither high nor low in M8 and A71 (mean M8= 6.328, mean A71 = 6.038). In A71 drivers scored between 1.6 to 10.83 and 2.16 to 11.5 in M8. Keeping in mind that NASA is 21 point scale, it is observed almost all scores fall in the lower quarter of the boxplot. Despite A71 and M8 being very different from each other there is no significant difference in workload. This can be explained with the help of malleable attentional resource theory (MART) as given by Young & Stanton, 2002. "Attentional capacity can change size in response to changes in task demands" (Young & Stanton, 2002, P.32); Environment across A71 and

M8 have changed considerably but attentional resources in M8 have shrunk to match task demands.

Table 6.1: NASA TLX Statistics

NASA Subscales	t-value	Effect Size (r)	df	p-value
Effort	0.46	0.03	65.99	0.64
Frustration	1.28	0.01	61.51	0.20
Mental Demand	0.16	0.10	65.99	0.86
Physical Demand	0.04	0.11	63.09	0.96
Temporal Demand	1.01	0.08	65.93	0.31
Performance	0.44	0.08	65.38	0.66
NASA TLX	0.59	0.06	65.96	0.55

No significant difference in workloads between M8 and A71 were detected (t (65.96) = 0.59, p=0.556). Effect size is the "degree to which the phenomena are present in the population" (Cohen, 1988, p.9). Table 6.1 shows barely detectable effect sizes across all subscales and NASA TLX. This means participants gave a verbal protocol and negotiated the test route almost effortlessly or with very little effort. In VPA, out of the 84 core concepts generated, all 84 of them concerned driver's external environment and fellow traffic user. Despite the difference being insignificant between the two road types the boxplot of A71 is longer than M8. This suggests that A71 had larger range of scores indicating more heterogeneous nature of road.

Null hypothesis for NASA TLX is proven true. That is to say, that I.V and D.V are not related to each other. The experiment had enough power (80%) to detect an effect size as small as 1.103 (~1) or greater but it didn't thus proving the null hypothesis. The effect that workload had on drivers of two road types are too little to be taken into design considerations. The low workload can be attributed to the control measures used.

Traffic demands like merging in heavy traffic, giving lane, etc. are found to have increased workload (Schneider et al., 1984 as cited in Ward, D., 1996). All the drivers were asked to drive in off peak conditions with almost no traffic thus having low workload.

According to DeWaard, 1996 slower driving speed is individual adaptation to negotiate task demands. From Figure 6.3: speed profile it is clear that driver drove slower than the posted speed limits. Hence, there is a possibility of driver to experience higher workload in A road than Motorway in peak hours. In this case, drivers masked their workload by driving slow. The difference in actual and posted speed limit is higher in A road meaning higher workload in A road. Looking at rear view mirror in unfamiliar environment is also attributed to be an overt sign of high workload by DeWard, 1996. "Behind" has occurred in 128 times in A71 and 109 times in M8 in verbal protocol analysis.

6.5 Self-Reported Measure

6.5.1 SART Results
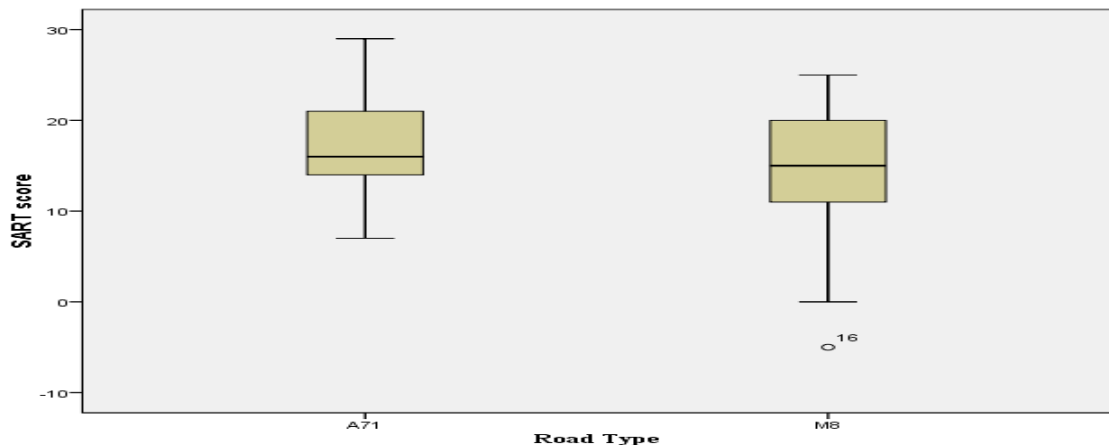
Figure 6.6 Boxplot of Situation Awareness rating Technique (SART).
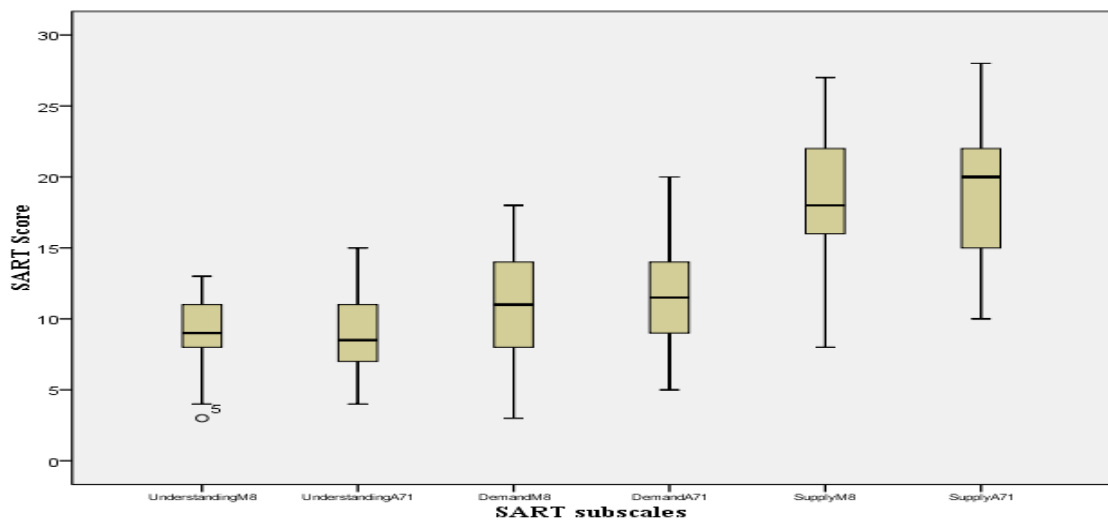
Figure 6.7 Boxplots of Subscales of SART

Situation awareness rating technique or SART is a very popular SA assessment technique. Given by Taylor in 1990 it is often used in conjunction with SAGAT. SART assesses the SA of an individual based on his opinions about the experience. The subject is asked to rate his response on 3 factors namely; understanding, demand and supply on a bipolar scale (Low = 1 to high = 7). The scales are then combined to give an overall score for any given system. S.A = U- (D-S). In this experiment, the 10 item (10D) SART was used. SART comes in different forms like 3D, 10D, 14D.

SART is constructed in the context of aviation sector. It is based on the definition, "situation awareness is the knowledge, cognition and anticipation of events, factors and variables affecting the safe, expedient and effective conduct of the mission" (Taylor, 1990, p.53). 84 Royal Air force personnel's were interviewed in 3 stages by psychologists. These interviews were semi-structured in nature based on repertory grid technique[1] and followed a fixed protocol for knowledge elicitation. The first stage was scenario generation which provided 43 scenarios. These were further reduced to 29 scenarios (18 high SA and 11 low SA). The second stage was construct elicitation.

15 personnel's participated in the second stage. Each construct was broken down by repertory grid triad method. 44 constructs were obtained from 29 scenarios. These

---

[1] Repertory Grid Method: Given by Keith Goffin, 1990.Help the interviewees to express their views on a complex task without interviewer bias. The repertory grid is the matrix of quantitative data.

constructs were subjected to statistical analysis and 10 generic constructs were obtained. Based on knowledge of attention and perception and for theoretical consistency Taylor further grouped these constructs under 3 broad categories (Demands on Attentional Resources, Supply of Attentional Resources and Understanding). Constructs which showed high inter correlation and loading were put in one category. So, instability, complexity and variability were grouped under demand; arousal, concentration, division of attention and spare mental capacity were grouped under supply; information quality, quantity and familiarity in understanding.

The main advantage of SART like all other self-rating techniques is that it is easy to administer, quick to score and non-intrusive. SART lacks any task specific question this makes it easier to apply in all possible domains. Although SART is subjective in nature but it gives an idea of cognitive processes involved in the process (Taylor & Selcon, 1991). SART may not provide the actual SA but it gives an indication of operator's level of confidence regarding his SA (Endsley, 2000). SART is found to be sensitive to design variables in a system (Selcon & Taylor, 1990). SART dimensions were derived from aircrew knowledge and experience; this makes SART high on ecological validity (Selcon & Taylor, 1994, Stanton et al., 2004). Demand group of constructs are found to have high sensitivity towards skill based tasks and the number of tasks subjects were asked to perform (Selcon & Taylor, 1994). This means SART can distinguish between workload and SA (Vidulich, 1992). Constructs in understanding and supply were found to be sensitive to cognitive (rule based tasks) and continuous tasks respectively (Selcon & Taylor, 1991). It is unobtrusive in nature and has high face, predictive, content and concurrent validity. Its construct validity is open to debate (Uhlarik & Comerford, 2002).

In this experiment SART was administered to participants twice. After finishing the drive on A71, A899 drivers were asked to come to a halt at BP Petrol Station. They filled out SART and started their drive on M8. On reaching Heriot-Watt which is at the end of M8 drivers were asked to complete SART this time emphasizing on their experience at M8

An independent sample t-test was administered to test the effect of road types on subjective situation awareness, t (58.32) =0.027, p=0.081, higher scores of SART was

found in A road while lower scores were obtained in M8 (SART A71= 17.352, SART M8= 14.352). The experiment was powerful (80%) to detect an effect size as small as 2.972 points on SART scale. Thus, I.V has a very strong effect on D.V. Null hypothesis is rejected, variables are related to each other (table 6.2).

Table 6.2: Statistics of SART Subscales

SART Subscales	t value	Effect Size (r)	df	p value
Understand	0.464	0.057	65.996	0.472
Demand	0.809	0.099	65.971	0.748
Supply	0.527	0.065	65.143	0.752
SART	0.027*	0.003	58.322	.081**

significant at 5% level

**significant at 10% level*

SART scores are found to be higher on A71, A899 and M8. Caution must be exercised in interpreting these scores in light of DSA. High score here means the driver perceived his SA to be higher on A71, A899. This doesn't mean that the SART score is his/her actual SA. According to DSA model, road environment is viewed as a system and SA is the property of that system. If a driver here has higher SA this means that the road environment has lower SA and vice versa. To drive safely the driver had to develop higher SA to compensate for low SA of A71. This is revealed in VPA.

Following DSA model, SA of a system is maintained by SA transactions. Since the inanimate part of the A71 road environment had lower SA higher SA transactions has taken place from the driver. The driver thus perceived successful driving performance in a bendy curvy A road junction to be SA powered activity thus the higher scores on A71. Similarly, low scores on M8 indicate towards the SER nature of motorway as proven in the pilot study. The road has high SA in itself which didn't require high amount of SA from the driver. So driving on a steady 2 lane with less variable speed limit motorway was regarded by drivers as fairly straightforward.

Figure 6.6 shows there is more variation in the SART scores of M8 than A71, A899. Almost equal number of subjects has scored higher and lower than the mean in M8. The scores of M8 are in agreement with each other. In case of A71 most of the scores lie above the mean. This means most subjects have scored higher than 17.352. This again reinstates that due to lower SA of the road the driver had to exercise more SA of his own to keep the SA of the system intact. It is interesting to note the boxplot of supply of attentional resources in Figure 6.7. Clearly, the boxplot of M8 has shrunk to adjust with the low attentional demands on the road. Also, supply boxplot of M8 has higher spare mental capacity than A71. Obviously, if the maximum capacity of the sample has been reached as result of the task like the one in A71, they cannot perform successfully in an event of emergency (Young & Stanton, 2002).

Similar trend has been observed in boxplots of understanding and demand. Especially in understanding the boxplot of M8 is smaller than A71, A899. This implies greater distribution of scores on understanding in A71. This means some drivers could comprehend more and some less. On the other hand in case of smaller boxplot of M8 drivers have more or less understood similarly. This points out towards the homogeneous nature of M8 like any other motorway. The Demand boxplot for M8 is very close to a symmetrical plot whereas that of A71 is slightly positively skewed. Additionally boxplot of M8 is longer than that of A71 again indicating towards a greater variation of scores. In summary, according to SART both the roads almost impose similar demands on the drivers. But the design of the roads are such that M8 facilitates in comprehension and understanding of those demands better than A71 with good supply of attentional resources. Furthermore these SART scores prove that SA is imbedded in the system through design thus invoking lesser SA from drivers.

6.5 Conclusions

The sample chosen is very well aligned with the wider population of British drivers. No driver with inherently dangerous driving style was found through DSQ.

The qualitative results proved that 35 UK drivers drove 22.5 miles long route with very little workload. They interpreted their performance to be better in A71 as the road

design is more challenging than M8. These results reinstate the results obtained in the pilot study which is exploratory in nature.

Drivers have driven 11.94 mph ~12mph lower than 50 mph which is the average speed of A71. In M8 this discrepancy is 6.6 mph lower than posted speed limit. This suggests M8 is designed in accordance (or close) to SER principles which match with driver expectancies from the road.

The boxplot of NASA TLX A71 is longer than M8 indicating a bigger variance in workload scores. This implies that A71 is much more variable and unpredictable than M8. NASA TLX scores are insignificant across two road types. This means drivers drove with very little workload. Frustration scores are significant on M8. This is due to presence of lorries and uncontrolled environmental factors.

SART scores are found to be significantly higher on A71 than M8. Subjects perceived their SA to be higher on A71 by 3 points than M8. SA of a system is maintained by SA transactions. According to DSA drivers are also an agent of the road system like infrastructure and design elements on the road. The SA of a system will remain the same with varied distribution of SA across the agents in the road. So in A71, drivers increased their SA to maintain the SA of the system as the road environment had low SA due to poor design. This is not the case with M8.

The succeeding section will present the results of VPA in detail and discuss its implications.

Chapter 7 : Extracting Endemic Road Features

7.1 Introduction

The results of on-road study are discussed in Chapter 6 and 7.

While Chapter 6 discusses the results from control measures and self-reported SA, Chapter 7 on the other hand disseminates network analysis performed on verbal protocol obtained from the participants recruited in this experiment. Chapter 7 attains the fourth objective which is to extract cognitively salient 'endemic features' required by SER approaches using DSA methods and validate them. The first objective was attained by selecting distributed situation awareness (DSA) model to explain SA in road transportation domain (Chapter 2). The second objective in Chapter 4 proved that driver SA is composed of elements external to the driver. The third objective proved that there is a link between SA and SER. This was done through an on road exploratory study (Chapter 4). The outcomes of the objectives in the thesis contribute towards the foundation of user centred roads.

Results discussed in this chapter were obtained from naturalistic study. Although Chapter 3 and 5 explains what naturalistic studies are, a short summary here will be worthwhile. Naturalistic studies represent the most advanced form of data collection for transportation studies. It is recently developed as a research method to observe road users everyday driving behaviours. The observations take place in drivers own car. There are three reasons for using naturalistic driving in this thesis. First naturalistic study works within the interactionist perspective i.e. driver, vehicle and road are connected to form the traffic system together with other traffic and external factors (SWOV, 2012). It conforms to the DSA model for road transport and is therefore appropriate for the aim of this PhD. Secondly; naturalistic study provides experimenter with a range of unobtrusive data collection devices regarding vehicle and driver behaviour. Finally, naturalistic study has been used in large scale projects such as 100 car naturalistic study to study distraction in driving, 'Interaction' which explains why, how and when drivers use intelligent technologies in their vehicle and 'Prologue' which explores the feasibility of naturalistic driving study in Europe (SWOV, 2012). Thus

NDS are sensitive to driver behaviour change with its validity established by results from numerous other driving studies.

Results in this chapter discuss the results of verbal protocol method in terms of network content and network structure. Furthermore it also throws light on common and unique concepts derived from these networks. These networks are made from sophisticated software called Leximancer™. Details of Leximancer™ are mentioned in chapter 4 but biggest advantage of Leximancer™ is the complete repeatability of results. This means high reliability. Further reliability of Leximancer™ in driving has been established by successful use of Leximancer™ by Walker et al., (2009, 2012); Salmon et al., (2009); Salmon et al., (2010, 2013). Following network analysis, social network analysis was performed to determine the sociometric status of each concept. This was done with AGNA. Like Leximancer™, AGNA has also been validated in driving context and has complete repeatability.

7.2 Verbal Protocol Method

The verbal transcripts obtained were transcribed in Notepad and subjected to semantic network analysis. One of the verbal transcripts had to be discarded because of the poor condition of the vehicle. Thus, 34 network analyses were performed on 34 transcripts. Leximancer™ was used for network analysis.

A measure of information in the environment is provided by word counts obtained from verbal transcripts (Salmon et al., 2011). 82,594 words were obtained from the total route. Assuming the null hypothesis to be true 41,297 words should be generated from each road type. That would mean 3470.336 words / mile for A71 and 3895.943 words / mile in M8. In this experiment A71 generated 49,934 words i.e., 4196.134 word/ mile and M8 produced 32,637 words i.e., 3078.962 word / mile.

Table 7.1 Descriptive Statistics of Word Count

Road	Number	Mean	Std. Dev.	Min	Max.
M8	34	960.9706	356.250	836.6690	1085.272
A71	34	1468.6471	501.016	1293.8342	1643.460

The difference in the word count between two roads are significant (t (33) = 4.815, p = 0.000). This means word counts have changed with road type. This difference can be attributed to distinct design of the two roads. Salmon (2011) suggested that more challenging road types produce more words. This is illustrated in this experiment. A71 being the mixture of rural, built area and urban roads have produce 21 % more words than M8. Thus, drivers found driving on A71 to be more challenging than M8. M8 on the other hand, is purposely designed to help drivers make sense of the situation readily. Although drivers during the study have complained M8 to be bland and reported A71 to be interesting, but the higher number of words show that actually drivers have higher mental compatibility with M8 than A71. This is again in conjunction with the findings of pilot study which led to select M8 as the control condition.
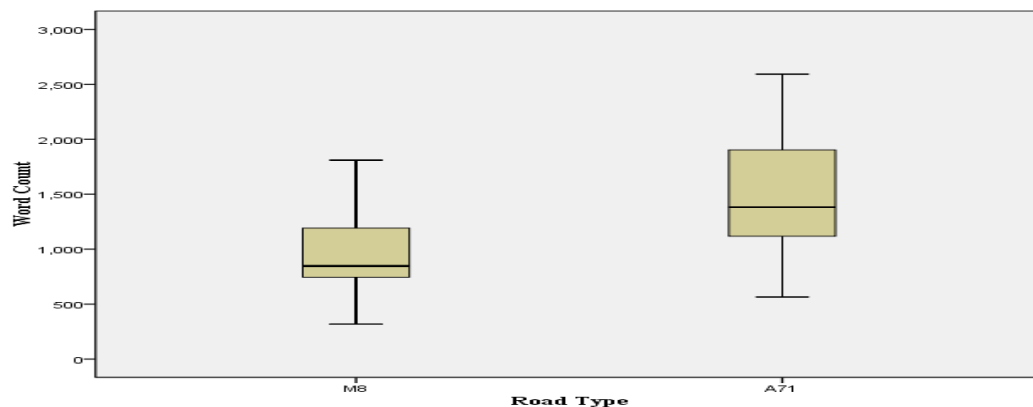
Figure 7.1 Box plot of word count across A71 and M8

From Figure 7.1 it is clear that A71 produced more words in total and per mile than M8. As mentioned before A71 has 21 % more words than M8. Figure 7.1 show that A71 has a larger range of words than M8. This indicates some drivers are extracting less information on A71 than M8 but this difference is not significant to have a design implication on a large scale. Secondly, this can also be attributed due to the uncontrollable traffic conditions.

The number of words is found to be directly proportional to the amount of semantic information extracted which in turn is consistent with speed and hazard incident rate (Walker et al., 2013). Drivers on A71 expressed more words; this means participants processed more information than M8. Higher number of words on A71 doesn't mean

higher SA on A71. It actually means participants processed more information on A71 but not all of them contribute to driver SA. More number of words can also be attributed to poor design of A71. Since participants are not getting the required information on the road readily so they process a lot more elements then needed and make sense of it to drive the vehicle forward

Speed and hazards are important elements of road design. As mentioned above these design elements are steady with the amount of semantic information. Thus, A71 being rich in number of words per mile is likely to be more complex in terms of road design. Hazards are defined as, "anything dangerous or potential to cause the driver to change the speed of the vehicle" (Coyne, 2007). A71 has more variable speed limits and higher number of hazards ("Bends need to get a lot better here." "There is a sign of camera but I don't know if it is a speed camera." "Road work signs"). Hazard perception is called SA in dangerous situation (Horswill & McKenna, 2004). Higher number of hazards is likely to hinder hazard perception which will lead to poor SA. Hence, it is fair to say that on the basis of verbal protocol A71 is found to have low SA. Furthermore, number of words is indicative of the presence of number of stimuli (Walker et al., 2011). Clearly, A71 has higher number of stimuli present, all of which are not appropriately designed and placed thus causing the driver to have lower situation awareness.

M8 on the other hand has essential recognisability characteristics (ERC as given by Crow, 2004). ERC are referred to as characteristics visible to drivers which can provide them cues on the road they are driving on (Stelling-Konczak et al., 2011). They are the initial building blocks of making a road self-explaining in nature. Some ERC which can be found in M8 are 'road curvature', 'smooth surfaces' and a fairly uniform road environment. The cornerstones of SER are homogeneity recognisability and predictability (Charlton et al., 2010). M8 is consistent with the design of other motorways ("like most motorways I have driven on", "this is like M11 having two sets of lane.") such as M9. It is also very predictable in terms of speed, road user types and manoeuvres. For instance, the speed limit for the whole 11.9 miles test route was 70 mph, road users were mainly cars, vans and Lorries. No cyclists or motorbikes were found or seen on the road. Driving on M8 is dependent on other traffic as the layout was very constant with two lanes. One is the overtaking lane and other being the outside lane for slow moving vehicles.

Another ERC of SER is recognisability (Stelling-Konczak et al., 2011). Due to the distinct road layout M8 it is easily distinguishable ("fantastic view"). In this experiment a stretch of A71 called A899 is found to have very poor recognisability. A899 is a dual carriageway road with 50 mph. Being a dual carriageway, drivers expected a higher speed limit ("I am surprised the speed limit here is 50 mph") but this did not match with the posted speed limit. Hence, A899 is not designed to match driver expectations and there is a conflict of schemas. SER researchers state that variation between roads of same category has to be less whilst a large difference is required to distinguish different types of roads (Stelling-Konczak et al., 2011). A899 was closer to M8 with respect to design but had the same speed limit as A71. This led many drivers travelling on that road to break the speed limits ("So all moving faster than 50 mph"). Ideally, A899 should be having a speed limit of 70 mph with frequent speed drops near the slip roads. Alternatively, a complete reconstruction of the road is required to make it single lane and giving wider space for slip roads joining in.

Results of VPA are very well aligned with the qualitative measures such as NASA TLX and SART as discussed in Chapter 6. The workload results prove that drivers did not have any effect of workload on driving. It should be noted that if the difference in workloads between two roads would have been significant enough to have a design implication, then they should produce less semantic content on A71. So workload was a controlled factor. It was a very effortless drive for the drivers both new and old drivers.

Chapter 6 show that drivers have regarded their SA to be significantly higher on A71 than M8. Verbal protocol result contradicts this. Walker et al., (2013) reported that motorways are richer in SA than A71 and they follow a SER design approach. This discrepancy in self- reported SA can be attributed to higher cognitive load imposed by A71 on the drivers. Cognitive load disrupts higher level actions by presenting too much apparently relevant information in a channel (Pass et al., 2003). Clearly, A71 is presenting drivers more stimuli to process therefore drivers are under the impression that they are maintaining a high SA. Design of A71 is such that it presents drivers with more stimuli which need immediate processing. Therefore, it disrupts higher process such as planning and guidance. Another reason for having a high word count and perceived higher SA in A71 can be understood with the help of Neisser's perceptual cycle (1976). Through guided exploration drivers sample the environment which

modifies driver's knowledge. This cycle is not able to take place in A71 due to its variable nature. It changes frequently and is inconsistent with other roads of the same category. There are frequent road blocks such as road works, patchworks on road surfaces, sharp bends and conflicting road signs which hinder this cycle to complete.

These results of word count also support the score of individual SART subscales by the drivers. Drivers have perceived their SA in A71 to be higher than M8. This means A71 to drivers was more demanding than M8 (+0.206). Thus again supporting the low SER nature and poor SA of A71. Drivers have scored higher in understanding in M8 (+0.353) than A71. Higher understanding score in M8 can be attributed to designing road with higher ERC and its design to be closer to that of SER. Furthermore, drivers have scored +0.529 points higher in supply of resources in A71 than M8. Like SART scores with more words don't mean more SA. It implies more semantic information is extracted from A71 rather being presented by the road to its drivers. In A71 drivers had to extract S.A concepts from the road due to numerous dynamic stimuli. SART scores reveal that drivers have perceived this extraction process to be their SA. Thus, it is assumed that drivers will perceive their SA to higher on rural roads, country road and so on as drivers correlate it with their performance thereby generating more words.

7.3 Knowledge Networks

7.3.1 Propositional Networks

This chapter talks about extracting endemic features and this will be done through propositional networks. They are well established form of knowledge representation (Walker et al., 2010, 2012) "Propositional networks are networks depicting the information underlying a system's awareness and the relationship between the different pieces of information" (Salmon, 2008, p.64). Propositional networks represent DSA as knowledge objects in the network and the relationship between these concepts. (Salmon et al., 2009). This agrees with the underlining concept that the DSA of system is not simply arithmetic addition of SA of different agents in a team working in a complex environment, but a property of a system which emerges from the interaction of different information elements or agents making up the system. Propositional networks also

negate the idea of "more the merrier" in a complex system. 10 elements in a network strategically connected to each other will not result in higher SA of a system than having only 2. It is the intricate cause and effect relationship between the elements that makes one element more important than the other (Walker et al., 2010), but all of them work together to contribute to the SA of the system.

Propositional networks are not a novel concept. The term was first used by Quillian in 1968 although Anderson and Bower (1973) argue it to be as old as Aristotle. Semantic networks represent the complex structure of human knowledge (Trehub, 1991). Every concept gains its meaning in terms of its position in the web of relationships (Trehub, 1991). The nodes in the network denote the knowledge objects or word concepts. The links denote the connectivity between the concepts. Trehub, 1991 defined semantic networks as , "propositional knowledge network structure consisting of a set of nodes that are selectively connected to each other by links labelled by the relationship between each pair of connected nodes" (p.99).

Ogden (1987) stated definition of any situation can be presented through semantic networks. This has been put to practice in on-road study conducted. As explained in previous chapters drivers produced a verbal protocol in a driving situation. This was transcribed to produce semantic networks with the help of Leximancer™. The software extracted the noun like information elements from the transcripts of the drivers to form nodes. The verb like links was formed between the nodes. The aim of this was to determine the structure of driver SA. Length of the links represents the strength of the relationship between the concepts. Longer links mean weaker connection and smaller links represent strong connection. According to DSA, driver SA should consist of outputs of all agents in the road environment like, crossing, junctions, trees, pedestrians thus proving again that SA is spread across the environment. Consequently, to help the driver drive safely an attempt should be made to make the semantic networks simpler by altering the road design.

In this experiment 35 participants provided verbal protocols whilst they drove the 22.5 miles long route. One was rejected due to poor quality of the vehicle. Therefore, 34 protocols were obtained and analysed. These protocols were transcribed in a word processor. Each transcript was fed into Leximancer™. Leximancer™ is defined as, "a

software system for performing conceptual analysis of text data in a largely language independent manner. The system is modelled on Content Analysis and provides unsupervised and supervised analysis using seeded concept classifiers" (Smith, 2003 p.23). Leximancer™ creates the semantic networks in 6 steps.

1. Conversion of raw text data (definition of sentence and paragraph, boundaries, etc.).

2. Automatic concept identification (keyword extraction based on proximity, frequency and other grammatical parameters).

3. Thesaurus learning (the extent to which collections of concepts 'travel together' through the text is quantified and clusters formed).

4. Concept location (blocks of text are tagged with the names of concepts which they may contain).

5. Mapping (a visual representation of the semantic network is produced showing how concepts link to each other).

6. Network analysis (this stage is not a part of the Leximancer™ package but was carried out as an additional step to define the structural properties of the semantic networks).

From the Leximancer™ analysis 82 concepts /knowledge objects were extracted for A71 and M8. Out of them 43 (52.439%) of belong to A71 and 39 (47.560%) to M8. This is consistent with the results of VPA. From VPA it is concluded that poor SA will lead to longer verbal transcripts and more words. This consequently produces more nodes. Furthermore, higher number of nodes is also consistent with the results of SART. Drivers perceived their SA to be higher on A71 where they had to engage in much more mental work and search for information to make sense of the environment. M8 as mentioned above has many essential recognisability characters, thus, producing lesser nodes. For example, a 68 years old driver on A71 said, "I don't know what they are doing. Roadwork's all over the place. They just spring up from nowhere." The same driver on M8 said, "sign on gantry saying roadwork's." Hence, the same driver understands the same situation in 2 different forms of environment very differently.
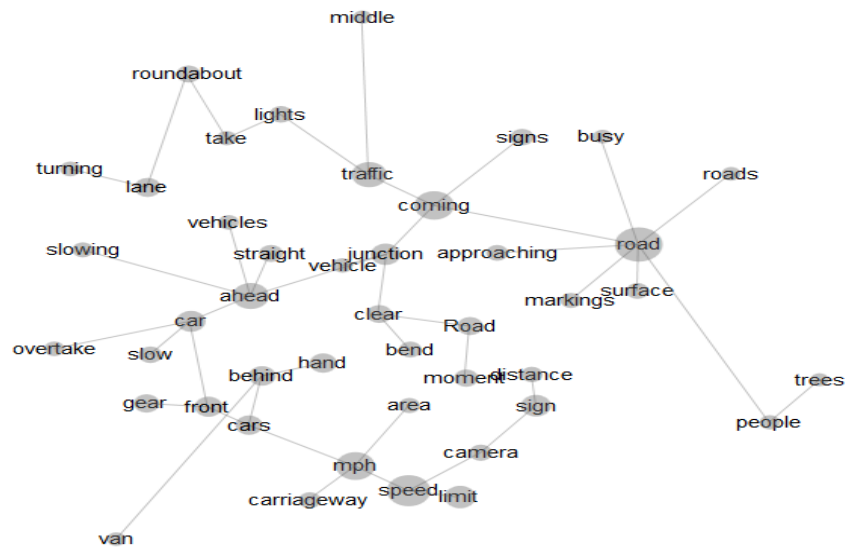
Figure 7.2 Leximancer™ output of A71

Figure 7.3 Leximancer™ Output of M8

The networks of M8 and A71 (Fig 7.2 & 7.3) differs markedly on standard deviation (SD) of centrality. S.D denotes how far a variable is from the mean. The S.D of A71 is 8.129 and that of M8 is 4.718. The S.D of A71 is almost double to that of M8. The centrality of M8 concepts are well aligned with the mean centrality of the network (21.968). A reasonable explanation for this is the good design of M8. The concepts which actually are the road elements are well designed and doing what they are supposed to do (serving their purpose, high functionality) in the road for forming the driver SA. So none of the concepts are too central and serving many purposes.

172

In A71 (mean= 24.33), the higher S.D indicates the non-uniformity of the concepts. This also indicates a mismatch between the design and how the drivers perceive it. In this situation some concepts will be used more than others. Also for those concepts that do not interact with other elements of the road the driver has to make up for it by putting in more mental effort or searching for relevant cues. For example, "junction" in A71 is connected to "vehicle", "coming", "clear". Driver SA around junction on A71 is basically made of direction and presence of other vehicles. However, junction in M8 is made of "traffic" and "lights". Driver SA for junction in M8 is connected to that of roundabout. The presence of roundabout activates the schema of junction ahead and this will be facilitated by the traffic lights on the roundabout.

7.3.2 Common Concepts
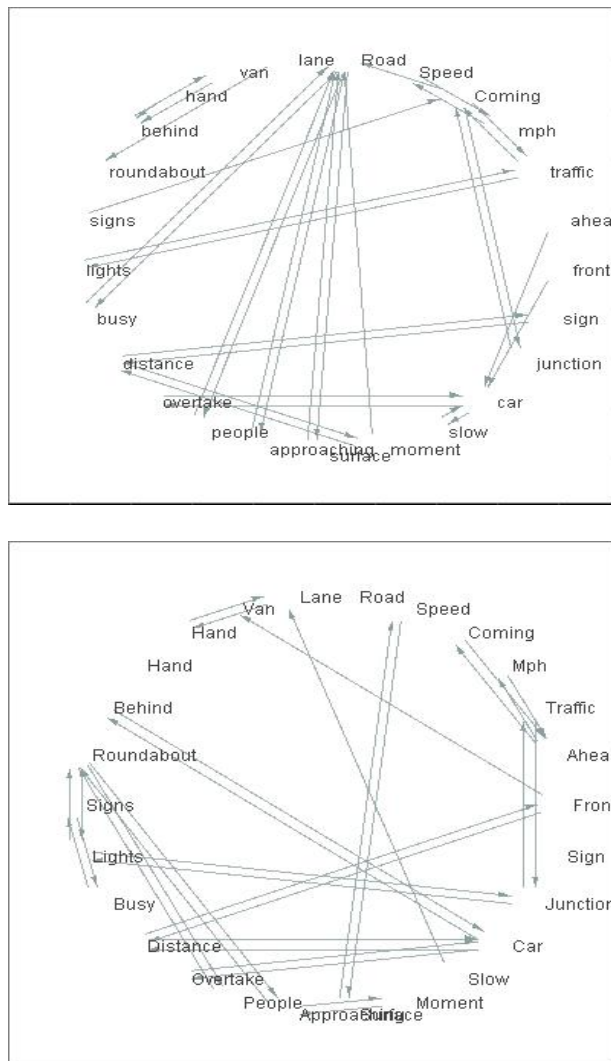
Figure 7.4 Social Network Analysis of A71 (top) and M8 (bottom)

Out of 82 nodes 25 knowledge objects are common across A71 and M8. So, in A71 out of 43, 25 concepts i.e. 58.193% are shared with that of 64.102% nodes in M8, which are 25 out of 39 concepts. Under null hypothesis the concepts will have the same inter-connectivity with the nodes in the network, but this is not the case. This is apparent from Figure 7.4.

The first difference amongst the two networks is with respect to diameter. A71 is slightly smaller (4) than M8 (5) in diameter. Diameter is the largest geodesic distance in the network (Hanneman & Riddle, 2005). So, in A71 no concept is more than 4 concepts apart from one another. This is a fairly dense network and concepts are close to each other. M8 on the other has slightly higher diameter of 5. Smaller diameter is attributed to higher road speed and presence of more hazards (Walker et al., 2011). This has been established before that A71 being poor in ERCs has variable speed limits and more hazards like "blind corners", faded markings", etc. Smaller diameter also indicates that drivers have direct access to semantic concepts (Walker et al., 2011). This kind of a network decreases the prominence of individual concepts (Walker et al., 2011) or in other words the endemic features. This reflects the presence of broader and in most cases not relevant concepts like, "pedestrians", "turning", "trees", "parked cars, and "badgers" on A71. M8 on the other hand do not have any such irrelevant concepts to focus on. Thus, attention on M8 is focussed on much reduced but nonetheless important subset of concepts (Walker et al., 2011). Additionally networks like M8 have more unique concepts.

Another difference between the common concepts is in terms of number of edges. The number of edges in A71 is 61 whereas that of M8 is 55. Since this difference is not very large the difference in density in the two networks is negligible (A71= 0.049; M8 = 0.055). Density indicates the level of interconnectedness between the concepts in a network. High density values suggest a presence of richer semantic links and well integrated concepts. Density in a network is directly proportional to the diameter of a network.

The most pronounced difference in the network is in centrality. The centrality of M8 is 5.247 and that of A71 is 15.744. 60% of the common concepts increase in centrality in A71, 12% remain unchanged and 28% decrease in M8 compared to A71. In summary

the number of shared concepts double thus showing an increase in centrality in A71, meaning they are semantically closer in A71 than M8. These differences in centrality and diameter suggest that despite the 25 concepts being identical in A71 and M8 they are markedly different from each other. To conclude, content does not determine or dictate structure. They are independent from each other.

7.3.3 Network Content

In this section we are going to see from a wider perspective (master network) into the common concepts which differ markedly (centrality wise) across A71 and M8. The largest difference in centrality is found in "moment" (+32.22 A71), "approaching" (+20.191 A71), "car" (+5.626 A71), "front" (+6.304 A71), "behind" (+5.631 A71), "lane" (+11.933 M8) and "traffic" (+7.173 M8). Rest of the 18 common concepts have a difference of 4 centrality points or less in M8 and A71.
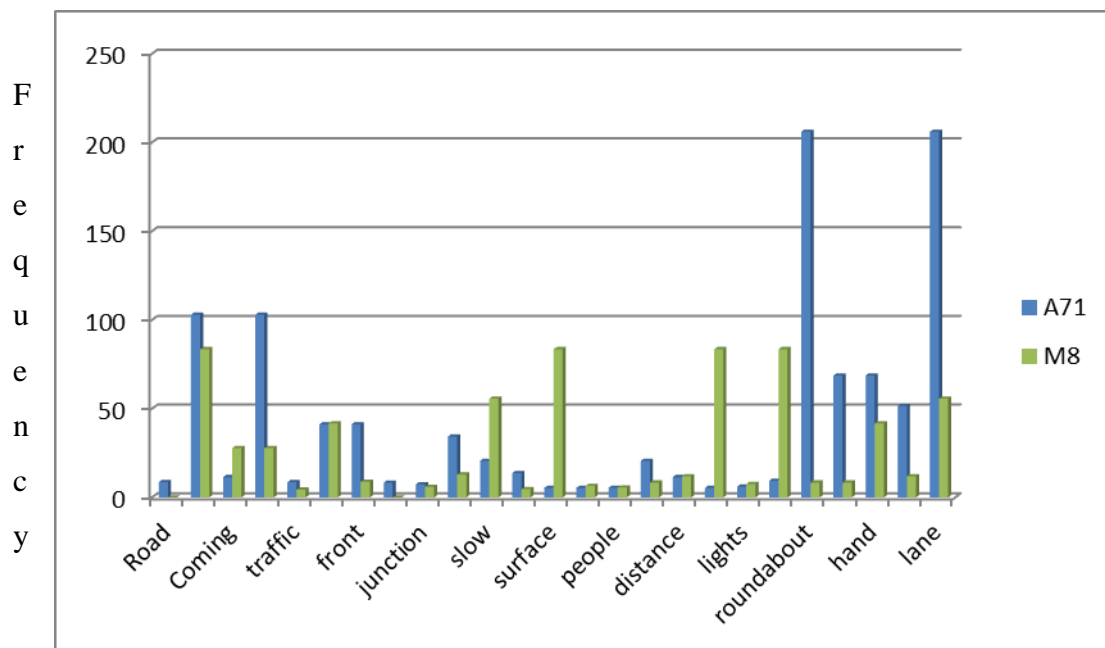
Figure 7.5 Distribution of common concepts across M8 and A71

Concepts common across M8 and A71 relate to "speed" (mph, sign, overtakes and slow). , the "road" environment itself (junction, roundabout, lane, lights, and signs) and "hand" (direction) of traffic mainly what is coming towards them (behind, front, ahead, coming, and approaching). Therefore driver SA is made up of the road environment and elements in it. This is in accordance to DSA and reinstates that this is the right and up to

date model to explain driver SA. 'Speed' is a very important concept for drivers on M8 and A71.

The tendency of drivers on both the road types is to know what lies ahead of them ("distance", "van", "car" and "moment"). They are very curious to know what is going on in the next 200 yards. This makes an interesting point with regards to how much information regarding distance should be provided to drivers in advance. Another conclusion to draw here is from the connectivity of these common concepts. For example, "speed" is connected to 4 concepts in M8 whereas it is connected to 7 in A71. This means drivers on M8 are interpreting the information of speed from 4 knowledge objects. They are, "surface", "overtaking, "doing" and "ahead". The same information in A71 is interpreted from 'speed limit' 'sign' and the 'area' in addition to the 4 concepts in M8. It is further controlled by the cameras installed which they just drove by and what type of vehicle in front of them.

"Surface" is a common concept but is connected very differently in M8 and A71. Yet another example where a same knowledge object is interpreted differently by same drivers on different roads. "Surface" in M8 is connected to speed. This means the good quality smooth surface is helping the drivers to maintain a speed of 70 mph on M8 ("running surface really good", "very wide road surface", "..very consistent quality of surface"). This also indicates towards the homogeneous nature of M8. On the other hand, "surface" in A71 is connected to "road" indicating the quality of road in general. For example, "change of road surface due to recent road works". "road surface still dry", "surface signs are very worn in this area." Hence, drivers are integrating this concept differently to make informed decisions.

"Busy" is found to be connected to "signs" in M8 and to "road in A71. First point to make here is M8 designs helps drivers to plan further through purposeful signage (such as gantry) by helping drivers choose an appropriate schema. A71 road design doesn't support this planning. Drivers don't get to know the traffic flow unless they are actually there. For example, "lots of junction here, quite busy traffic movements."

"Roundabout" occurs 128 times in A71 and 86 times in M8. The advantage of having roundabouts is it keeps the traffic flowing around a central island. Therefore, it requires

significant amount of interaction from other traffic to negotiate it successfully. In A71 roundabout is used for navigation. It is connected to "lane", "take", and "lights". Roundabouts in A71 are very well marked telling the drivers which lane to choose to arrive at a destination. The information for a roundabout coming up is communicated to drivers through signs. In M8 it is connected to "people", "approaching", "lights".

"Moment" represents the real time information for the driver e.g., "there is no one in front of me at the moment". This means the way it is connected to the other nodes will represent how the driver is deriving this information. In A71 moment is connected to "road" and "distance". If we zoom out then we see that the drivers become exceedingly concern about the real time information on approaching a junction. They update this from inbuilt road elements like traffic signs ("traffic lights are green at the moment"), geometric design of the road leading to the junction like "bend" and "straight" and traffic conditions on it like "clear". "Distance" on the other hand is feeding into moment to moment information of a speed camera coming up and road signs. For instance they used the information from the signs to predict the kind of road users ahead ("keep a safe distance here, kids emerging from the bus stop."). "Moment" in M8 is connected to "approaching". In the present study drivers on M8 have encountered 2 major roundabouts (Hermiston Roundabout and Livingston East Roundabout). So drivers have got the real time information and prepared themselves by taking appropriate lanes by getting the information of approaching roundabout from the road signs. From a careful observation of the master network of M8 it can be concluded that the moment to moment knowledge for the road for a M8 driver is to do with vehicles around them. They are very keen on maintaining a safe distance (headway) from the vehicles ahead of them. In fact it will be fair to say that to some extent that the speed of vehicles on M8 is dictated by vehicles in front of the. For example, regular users of M8 said, "I am going at 65 mph to give them a bit of distance." "Keep a nice safe breaking distance." On M8 drivers also seemed to be more anxious to overtake, so they tend to look for big gaps or distance between cars so that they can overtake and fill the gap.

"Approaching" is connected to "road" in A71 and "roundabout" in M8. Oxford dictionary defines approaching as "action of coming near/nearer to someone/something in distance and time." This means if driver is using the word approaching it implies that he has to make a decision under significant time pressure of what lies ahead of him.

Distance and time are factors contributing to SA (Endsley & Rodgers, 1994). The high occurrence of "approaching" in A71 (79) than M8 (36) is the indicator of poor SA in A71. For example, one of the drivers used "approaching" 12 times in A71 and not once in M8. Drivers on both the roads are on the lookout for events ahead of them. The drivers on A71 show an increased awareness of vehicles in front, behind and on side when near a roundabout and intersection. In A71 the driver's schema for "approaching" is oriented to traffic signs ahead, bends on left/right, cyclists behind, slip roads, staggered junctions and roundabouts coming up. That is mainly road elements. In M8 "approaching" is used with reference to end and beginning of junctions and the upcoming queues leading to a roundabout.

"Lane" (298) is the most central and highly occurring concept in M8. Lane is connected to "fast", "slow", "hand", "inside", "outside", "cars" in M8. In A71 it is connected to "turning" and "roundabout". It occurs 157 times, 29%. An interesting observation to be made here that lane, the most dominant concept in M8 is related to the road environment itself. This is consistent with the findings of Walker et al., 2013 stating that motorway represent SER in the U.K. Driver SA in M8 seems to be made of road environment itself. The idea of lane in A71 is activated on approaching a roundabout. Roundabouts like Oakbank and Lizzie-Brice have a selection of lanes. Drivers on M8 interpret events with reference to lanes configuration and shoulders. For example, "there is no one on the outside lane now". "Nice clearing on the outside lane now."

"Traffic" is an uncontrolled variable in this experiment. The general nature of motorway certainly led to high occurrence of "traffic" in M8. Traffic is strongly connected to "lights" in A71, indicating frequent occurrences of traffic lights in A71 ("traffic lights ahead"). Traffic is connected to "motorway" and "junction" in M8. This implies driver SA is not only concerned with the traffic in front or behind but also coming from opposite direction. Drivers on M8 are generally concerned with the volume of traffic at the road they are on ("traffic conditions as far as I can see are pretty quiet actually"). On the contrary, drivers on A71 are concerned with traffic that is in their lane or coming towards or behind them ("going through the roundabout, lots of traffic"). This would explain it being connected to "middle" and "coming".

"Distance" occurs 65 times in A71 and 52 times in M8. This implies drivers have given almost equal importance to "distance" in both road types but contribute differently in driver SA. In M8 drivers have used "distance" as valuable information to "overtake" a vehicle. Whereas "signs" in A71 communicate to the drivers the "distance" of speed cameras from their current location. Future technology should focus on presenting this information in a digital form. Interestingly, "distance" is not connected to "front" or "behind" in A71 unlike M8. This indicates drivers on M8 are not just overtaking a vehicle because it is slow or slowing down or big but they are making an intelligent choice to overtake it on the basis of distance between them. Thus, drivers are maintaining a safety margin on M8. "Front" and "Behind" are similarly connected concepts in A71 and M8. 28% of concepts in A71 relate to front and 29% in M8. Behind occurs slightly more in M8 (37%) than in A71 (24%). Drivers on both the roads look out for vehicles in front and behind of them. Higher occurrence of behind in M8 means drivers on M8 are more conscious what lies behind them which would mean more "mirror checking" in M8 than A71.

7.3.4 Unique Concepts/ Endemic Features
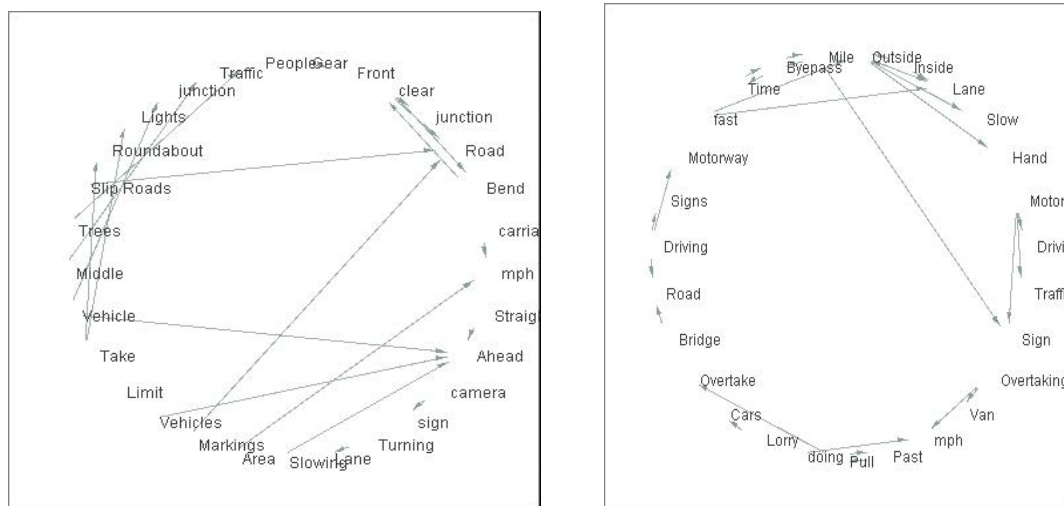
Figure 7.6 Unique Features in A71 (L) and M8(R)

Table 7.2 Cognitive Salient Features of A71 (L) and M8 (R)

Concepts	Centrality
Turning	13.188
Roundabout	16.682
Take	19.011
Middle	21.168
Roads (slip)	21.529
Vehicles	24.105
Vehicle	30.258
Clear	25.638
Trees	18.454
Camera	20.904
Limit	18.975
Carriageway	20.071
Area	20.071
Behind	21.809
Gear	21.079
Straight	50.633
Slowing	24.105
Markings	21.529
moment	48.21
approaching	38.311
Junction	30.441

Concepts	Centrality	
Outside	22.55602	
motorway	29.54348	
overtaking	25.53247	
past	13.25854	LOWEST
doing	22.41649	
inside	23.18124	
lorry	24.54176	
pull	14.69189	
bridge	18.4898	
driving	25.16667	
fast	22.79245	
time	15.80233	
mile	20.43609	
byepass	17.88158	
speed	22.60291	
slow	22.41649	
surface	19.37968	
busy	18.4898	
signs	21.40158	
coming	33.555	
cars	29.95041	
mph	26.844	
lane	26.64706	
ahead	29.95041	
traffic	32.74699	

Concepts marked in red are added from the master network.

Important differences in driver S.A of A71 and M8 are revealed through the unique concepts (Figure 7.6). Being unique would also mean they are endemic in nature i.e., native or restricted to a specific area. 20 (46.561%) endemic features of A71 and 27 (69.23%) of M8 were extracted. Clearly, SER nature of M8 has led to higher number of endemic features in M8. The advantages and effects of endemic features on driver behaviour are discussed in detail in previous sections. An endemic feature or a group of them will make a road easily recognizable and predictable. Easy recognition would generate appropriate driver schema. As activating wrong schema is one of the major causes of human error (Norman, 1981) so endemic features rich road will put this problem to rest. From the master network it is observed some the concepts have scored very high on centrality despite being common concept. Thus, the aim of this section is to derive a list of **cognitive salient features (CSF)** encompassing such concepts. All the unique features are endemic but not necessarily cognitively salient. CSF are those which are most important to the driver while driving on a particular kind of road. So, they have to be the right mix of unique and common concepts. Certainly, almost all or most of the unique features will be salient however; those (from master network) scoring above 1 SD from mean in A71 and M8 will also be considered as cognitive salient. Table 7.2 above represent the cognitively salient features in two road types. The ones in red are common across two road types but are far more prominent with respect to centrality in either A71 or M8.

"Straight", "moment", "approaching", "clear" and "vehicle" are high scoring cognitive salient features in A71. In M8 "coming", "cars", "lane", "mph", "junction", "ahead" and "traffic" are high scoring salient features. Thing to note here is some concepts / knowledge objects like "traffic" and "lane" are common across A71 and M8 so they are clearly not endemic. Endemic is "native or restricted to a certain place". They are viewed differently by drivers due to their design in different road types which makes them more or less important than the others.

Table 7.3 Descriptive Statistics of A71 M8 Centrality

Road Type	N	Mean	Range	S.D	Variance
A71	22	37.45	24.0466	9.716	94.408
M8	28	20.30	23.5444	5.260	27.669

Centrality values close to mean represent SER nature of a road. Concepts which are not further than 1 S.D from mean are regarded as cognitive salient. That is they are not very important and do not contribute to driver SA. If the concept is unique or endemic in nature but low on centrality; it will mean that knowledge is unique to that road but not actively participating SA transaction. Implying they are not contributing towards cognitive compatibility between road elements and drivers.

The number of cognitive salient features in A71 is found to be 22 which are 51.62% of total knowledge objects. 28 cognitive salient features were found in M8 which is 71.79% of total concepts (table 7.3). So, M8 20.17% has more CSF than A71. In M8 drivers are extracting less but meaningful semantic information. Whilst on A71 they are extracting more but only half of them are cognitively salient in nature unlike M8.

7.4 Cognitively Salient Features

"Coming" is cognitively salient feature in M8. This means that "coming" has also occurred in A71 but in a different context and does not contribute to driver SA as much it does in M8. In A71 "coming" is used to mention what is coming up with respect to road infrastructure like "bends", "signs", "junction", etc. "Coming" in M8 is used to describe "traffic" coming towards them. In some cases drivers have specifically mentioned the kind of vehicle coming from their sides or behind them. This reinstates that drivers on M8 are interacting their fellow traffic much more than what they were doing in A71. Whilst they interact on M8 the design of M8 is such that it provides the driver with the information the driver needs. From "coming" we can infer that drivers on A71 have to be much more aware. They have to know everything from the dynamic road design such as bumps, worn out markings, numerous intersections not to the steady flow of traffic. In A71 nothing really is constant or predictable. In M8 drivers are only concerned with the traffic overtaking them or who is behind them this is due to the reason that M8 is very predictable and in line with designs of other motorways.

"Straight" is the highest scoring concept in A71. It is ironic that M8 being an 11 mile stretch of straight long road has no occurrence of straight whatsoever. Straight in A71 is basically used to denote change ("so we are going to a twisty section having come along from long straight forward section") and planning (…"I have to continue on

A71….straight on"). This also implies drivers are interpreting the situation in terms of what lies ahead. Since driving on A71 is more to do with negotiating natural challenges like corners and bends; planning ahead becomes increasingly popular For instance, "straight across keeping on the A71. "Just approaching the roundabout for the Oakbank turning". In M8 information is provided instantaneously through gantries and suitable margins of error.

Some other cognitively salient features of M8 are "cars", "motorway", "driving", "overtaking" and "ahead". Ahead in M8 is used to understand traffic challenges, traffic arrangement, etc. ("there are cars ahead seems to be overtaking others", "several lorries up ahead") which is contrary to A71. "Driving" is an endemic and cognitive salient feature of M8. This has come up as the drivers were probed to say which road they preferred more for driving. The semantic network reveals that driving on M8 for drivers is a very holistic experience. It is connected to "road", "signs" and "motorway". It is interesting that how micro level analysis of M8 show that their SA is formed through interaction with other traffic but on a macro level they are viewing on M8 as a system of which they are a part of as well.

"Overtaking" here would be phenotype schemata. Phenotype schema or Local state specific routine is defined as "activated schemata brought to bear on a specific problem by a user" (Stanton et al., 2009). The idea of schemata has been described in detail in Chapter 2. In M8 the genotype schema would be going slow on the inside lane and faster on the outside lane thereby overtaking others. This implies that there should be a trigger (most likely for slow moving vehicles) of some sort which is making the driver switch to phenotype schema. The semantic network shows that overtaking is activated when there is a "van" in "front" or into some "distance"

"Motorway" in M8 is connected to "driving", "sign" and "traffic". This can mean a few things. First, is triggering of anticipatory schema. The end of M8 Junction 3 has traditional yard line signs indicating the end of M8. These "signs" triggers the schema of traffic condition in car drivers and likely presence of a "junction" ahead.

"Cars" is connected to "front", "behind" and "mph" in A71. Now "mph" has strongest connection with speed. This suggests that SA of drivers on A71 is made of cars coming

from behind as well as in front. Furthermore, they are concerned with the speed of the vehicles around them. In M8 "cars" have similar connection with the notable exception of "lane". This explains that drivers are actually anticipating the speed of vehicles not by mere guesses but by which lane the cars are in. Hence, reinstating and reproving the SER nature of M8.

7.5 Summary

Chapter 7 offers experimental evidence in support of SER-SA relationship. The results are analysed from an on-road study with 35 subjects and driving style and workload as control measure. The results demonstrated that number of words in A71 is more than M8. This is due to the fact that information/knowledge required for driving on A71 is not instantly available to drivers. Drivers on A71 had to process vast amount of irrelevant information and filter out the important ones. This hinders higher cognitive processes. In light of SA requirement analysis results this can be understood as road design of A71 is such that its core driving tasks poses high SA need from drivers thereby inhibiting SA transactions for global driving tasks.

Chapter 7 also showed marked difference in SA network in A71 and M8. Although over 50% of knowledge objects were found common across M8 and A71 but they are connected distinctly. This means 2 things first is knowledge objects cannot be studied in isolation. It has to be studied as whole within a network in light of its connectivity with other knowledge objects. Secondly, the results also showed that drivers on M8 are able to make an informed choice for overtaking. For example, drivers don't overtake on the basis of a vehicle's size or speed but assess the distance between the two in addition to other factors. Chapter 7 also demonstrated drivers are curious to know what lies ahead of them. This need needs to be supported by a SER design. For instance, in M8 "busy" is connected to "signs" whereas in A71 it is connected to "road".

The main conclusion of A71 is deriving a list of cognitive salient features. They are an upgrade of endemic features. Endemic features make a road what it is but cognitive salient features are those concepts which though active SA transactions form the perceived image of the road to a driver. So these features are a combination of common and unique concepts of a road. However, cognitive salient features are mainly composed of unique features which are also endemic but not all of them are active in driver SA.

Therefore, through social network analysis high centrality common concepts are included in the list of cognitive salient features. M8 was found to have 20% more cognitively salient features than A71. Deriving these features from propositional networks strongly establishes the link between SA and SER.

Since linking SER and SA is a very novel idea therefore, the following chapter will validate the extracted features on the observable driver behaviour.

Chapter 8 : Validating Distributed Situation Awareness Approach

8.1 Introduction

The aim of Chapter 8 is to assess if the extracted cognitive salient features in Chapter 7 are linked to actual driving behaviour. A secondary aim of this chapter is to re-establish the M8 as a self-explaining road in a classroom setting with different groups of subjects. Since the M8 has more cognitive salient features, it is assumed that the M8 will prove to be more self-explaining than the A71 in this experiment as well. This is done with a picture rating test. In this test subjects were asked to indicate the speed of choice in a piece of paper within five seconds of showing a picture of road. The insights of this validation study contribute toward attaining objective 5 i.e. develop a tool to assess the cognitive compatibility of roads and validate it against real-world accident data. The implications of Chapter 8 are discussed in Chapter 9 in the form of road drivability tool (RDT).

The purpose of this thesis is to lay the foundation for user-centred roads. A critical step in doing that is to establish a relationship between situation awareness (SA) and SER. Chapter 2 proved that SA is highly contentious lacking a universal definition and model. This is also supported by Salmon & Stanton (2013). Thus the DSA model is selected in Chapter 2 for use in examination of SA in road transportation. Chapter 4 explained the road environment as a source of driver SA. This is also supported by SER literature. According to these results, powerful changes in driver behaviour can be elicited through purposeful road design. This idea is then proved experimentally through an on-road exploratory study later in Chapter 4. The results were encouraging and supported expansion to a large-scale on-road study. A link was established between SER and SA by extracting endemic features through propositional networks. Through social network analysis a list of cognitive salient features were also obtained in Chapter 7. Due to the novel nature of these ideas and potential linkages, it is imperative to put these cognitive salient features to the test by assessing if these features can elicit any changes in driver behaviour.

Chapter 8 describes the picture-rating task (PRT) undertaken. It starts with describing speed as a behavioural measure in the form of reaction time, and how this relates to a

self-explaining road (SER) design approach. Furthermore, it explains the results of this experiment in terms of absolute speed and speed variance.

8.2 Background

The relationship between speed and road accidents is very clear and extensively investigated. For example, European Traffic Safety Council (ETSC) (1995) reported higher speed to be the core cause of road crashes as it reduces driver reaction time. Approximately 23% of accidents in UK are caused due to inappropriate driving speed (Sabey & Taylor, 1980). NHTSA (1991) has also stated high speed to be the cause of fatal accidents. Furthermore, the most common form of traffic offences is speed violations (Martens et al., 1997). Therefore, speed is key driver behaviour to test with cognitive salient features.

Driver speed choice is affected by the road environment, surrounding traffic, and the drivers themselves (ETSC, 1995). Edquist et al. (2009) identified road geometry (surface, width, geometry, grade, curvature, and delineation), roadside environment, road signs, markings, and self-explaining design as factors affecting driver speed. Edquist et al. (2009) also noted relatively uncontrollable factors such as parked cars, weather conditions, road works and other temporary factors influencing speed choice. Workload, fatigue, age, experience, etc. are some other factors which are also found to affect drivers' adopted speed (Edquist et al., 2009). Road surface is deemed to be the most important factor for speed choice (Edquist et al., 2009). Chapter 4 has proved experimentally and through a desktop study road environment to be the main source of driver SA, therefore it is essential to determine the effect the extracted road elements (which are the cognitive salient features in Chapter 7) on driver speed choice. This approach of determining the effect of the road environment on driver behaviour is consistent with the SER approach to road design.

Driver speed preference is affected by the perception of speed within the road environment, and the knowledge of the own speed. The latter is provided explicitly by a speedometer or vehicle- activated sign (Edquist et al., 2009). The former needs to be supported by the road environment. Road geometry and other characteristics will influence drivers' expectations of the appropriate speed for a road (Weller et al, 2008).

Therefore, it should be possible to influence driver speed choice through changes in road design, rather than merely relying on signage. The concept of user-centred roads involves designing a road system in which the driver's expectations (created by the road environment) are implicitly in line with safe driving behaviour which is appropriate for the road.

A picture-rating task is chosen for three reasons. Firstly, as mentioned above there is an intricate relationship between the road environment and speed. Therefore, showing different pictures of road and asking subjects to indicate their preferred speed is a logical choice. Secondly, speed perception from pictures is based on the paradigm of a picture-sorting task (PCT). PCT is found to be sensitive to detecting speed choices in experiments of SER (e.g. Kaptein, 1998; Kaptein & Classens, 2002). Finally, speed perception is a very practical measure, as it is easy to analyse and requires minimal time and resources. With appropriate road layout to evoke correct expectations from the driver, the largest reductions in speed can be seen and potential errors can be reduced (Martens et al., 1997). Hence, roads with CSF can change drivers' perceived ability to adhere to the posted speed limits.

The present experiments investigated to what extent the cognitive salient features determine behaviour. To that end, how speed choice differs in road environments in relation to the rise and fall of cognitive features was investigated.

8.3 Methodology

8.3.1 Research Design

The picture-rating task was administered in a within subject design. A 'within subject' design involves the same group of subjects in two different conditions. In this experiment the same group of people were shown all four groups of pictures. Out of a total 123 pictures from the total route, 70 were of A71 and 53 were of the M8. CSF are those handful of characteristics of a road which, through active SA transactions and compatible SA, form the perceived image of the road to the driver. Thus, determining these cognitive salient features is pivotal to road safety and any changes to road design should take this under consideration. CSF extracted in the preceding chapter will serve

as an independent variable (IV) here. Preferred reported speed by the driver will act as the dependent variable (DV). Variables controlled are traffic and weather conditions.

It is estimated that subjects will choose a speed closer to the posted speed limit for pictures high on cognitive salient features. Additionally it is predicted that pictures with more endemic features can help reduce variability in speeds reported by the drivers. This will imply that there will be less difference between posted and actually chosen speeds on the M8. Clearly, the difference between the average and the posted speed limit will increase as the roads become less endemic in nature.

8.3.2 Participants

Twenty male and female participants (Mean age =23.5, S.D= 3.215) with a valid driving licence voluntarily participated in the study. The participants had a mean driving experience of 5 years (S.D = 1.27). Before commencing the experiment a formal ethical approval was sought from the university. All subjects participated voluntarily and an informed consent was received from each of them.

8.3.3 Materials Required

Materials required included 90 blank pages (88 to be used, 2 spare); pencil; computer; The Microsoft Office PowerPoint with road images (sample shown in Figure 8.1) and; 20 participants.

Figure 8.1. Sample Picture of Road Shown to Participants

8.4 Procedure

In terms of actual procedure, this experiment was conducted with close proximity to the Dutch studies (e.g. Kaptein et al., 2002). As a first step, a pictorial route survey was carried out by viewing the test route on Google Maps. Screenshots were taken of a variety of sections of the 22.5-mile long test route. This produced a total of 123 pictures. Inter-rater reliability was assessed for 10% of pictures of the A71 and M8 to determine consistency among raters. Raters agreed on 68% of cognitive salient features identified in the screenshots, refer to Figure 8.1 ($p<0.001$, Kappa = 0.68).

All pictures were taken from Google Maps in reasonably dry weather conditions from the perspective of the car driver both in height and position of the road. Most of the pictures included other road users on the road to portray a more natural environment however some pictures did not include other road users. This actually corresponded to the original road conditions which the drivers drove on. Following this, the centrality of each picture on the basis of cognitive salient features in them was calculated. Knowledge objects of each road type were extracted through network analysis by Leximancer™ in Chapter 7. The social network analysis filtered out the cognitive salient features from other extracted knowledge objects. The pictures were then arranged in an increasing order for each road category (A71, M8). Finally, four groups were arranged, each having ten pictures. Groups were classified as:

a) Endemic A71 (first 10 pictures of highest centrality);
b) Non-endemic A71 (10 pictures with least centrality);
c) Endemic M8 (highest centrality), and;
d) Non-endemic (least centrality).

8.4.1 Instructions

Instructions given are detailed below:

You are about to see pictures of roads on the big screen. All pictures are different. Imagine driving on that road and ask yourself how fast you would drive on that road. Do not mind the legal speed limit; just consider the speed you would be driving on

there. Every picture will be shown to you for 5 seconds. So as soon as the picture appears you must make up your mind and determine a speed. 4 sheets of paper are provided with numbers at the side on each of them (Figure 8.3). Write your response in front of the correct picture number. If you miss any picture then please wait till the end of the test. Consider every picture separately. Do not make a mutual comparison between the pictures. Do not think too long; just follow your first impression. The test is non-stop. Series of pictures will be shown without a break so if you have any questions ask me now.

Figure 8.2. Participants of the Picture-Rating Task

Figure 8.3. Response Sheet Provided to Participants

8.5 Results and Discussion

8.5.1 Reported Speed
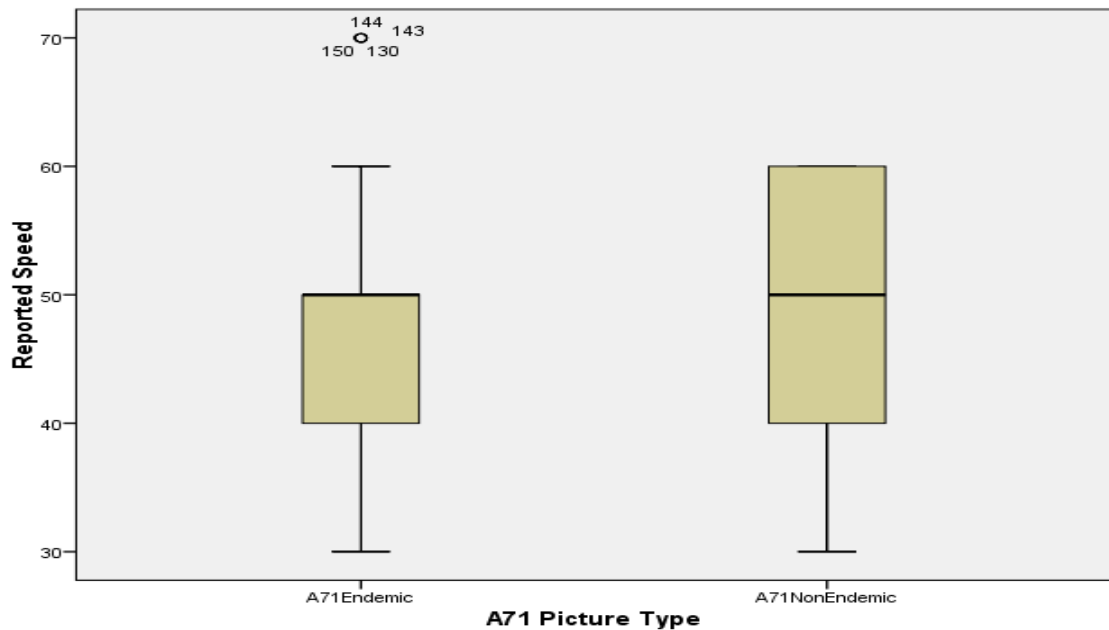
Figure 8.4 Box Plot of Differences between Absolute Posted Speeds and Absolute Reported Speeds on the A71

From the box plot of absolute speed on the A71 (shown in Figure 8.4) it can be seen that the range of scores is higher in endemic conditions than in non-endemic conditions. Descriptive statistics can be found in Table 8.1 below. Almost all subjects seemed to have scored above the mean for the A71 endemic category, whereas half of the participants have chosen speed above and below the mean speed of 50.25 mph. This suggests that each participant interpreted the pictures associated with the A71 non-endemic category differently from the other. This point towards less objectivity of road design of non-endemic sections of A71. The A71 endemic category on the other hand shows a steady trend of all participants scoring above the mean. This indicates that the road environment with high endemic or cognitively compatible features is interpreted in the same way by all car drivers. This level of objectivity can be attributed to affordances, gestalt principles, and some basic principles of good design in relation to human perception (Weller & Dietze, 2010). For example, the A71 non-endemic category has many fast transitions from A71 urban roads to residential areas, and then back to A71 rural. This makes drivers frequently switch mental models and activate required schema. Furthermore, these transitions are not marked clearly either by rumble

192

strip or speed bump. These road transitions however have been marked by road signs which can be often ignored or not be perceived due to psychological or physical filters (Weller & Dietze, 2010).

Table 8.1 Frequency Table of Absolute Speed

	A71		M8	
	Endemic	**Non-endemic**	**Endemic**	**Non-endemic**
N	200	200	200	200
Mean	48.1250	50.2500	67.5500	66.2000
Median	50.0000	50.0000	70.0000	70.0000
Standard Deviation	9.68651	10.04700	7086171	11.41101
Variance	93.829	100.942	61.807	130.211
Range	40.00	30.00	50.00	40.00

The mean speed adopted for the A71 endemic category is 48.125 mph whereas the posted speed of the A71 endemic category is 48 mph. Thus there is a mismatch in the driver speed expectation of only +0.125 mph. Participants on average exceeded the speed limit (negligibly) by +0.125 mph. Similarly, the mean adopted speed for the A71 non-endemic category is 50.2500 mph; however the posted speed limit of the same road condition is 47 mph. Hence, participants chose to drive in excess of 3.25 mph on the non-endemic portion of the A71. This mismatch can be attributed to poor SER design of the road. The road is not in line with driver expectations; therefore participants' perceived image is not as same as the objective state of the road.

Local roads such as the A71 and other B roads have been found to have the highest number of accidents per kilometre (Hamilton & Kennedy, 2005). Most of these accidents occur due to speeding, leading to head on collisions, running off of the road, and junction accidents (Hamilton & Kennedy, 2005). This is reflected in the higher speed choices of the A71 non-endemic category. According to Gugerty (1997) real-time tasks such as driving can be performed successfully by keeping track of a complex set of information, such as road conditions, own speed, other vehicle speed, location and many more. This is often not achieved in a non-endemic condition as there is less

number of cognitive salient features. This implies that there are lesser agents in the road environment holding compatible SA participating in SA transaction.

An interesting point can be observed in terms of standard deviation and variance between the two conditions of the A71. The variance of the A71 endemic category is 93.829 (S.D = 9.68 mph, mean = 48.12 mph) whereas the variance of the A71 non-endemic category is 100.942 (S.D =10.04 mph, mean = 50.25 mph). As pointed out previously, all participants indicated a similar choice of speeds in the A71 endemic condition. This implies that there is less variance in speeds of the drivers for the A71, due to the high number of endemic features. A higher number of endemic features will mean unique road user behaviour for a particular road. In this case, this means that the drivers on the A71 endemic condition were able to find the unique design elements (endemic features) which helped them to perceive the speed of the road required to drive safely. As mentioned previously, this perceived speed in an endemic road is closer to the posted speed. Thus the A71 endemic condition is able to produce homogeneous behaviour among its car drivers, unlike the A71 non-endemic condition.

A minor finding, or rather an observation, is that more road design elements do not necessarily produce a self-explaining road. This has been mentioned in Chapter 6. This supports that fewer features which have been strategically placed – i.e. designed ergonomically to fit within the road environment, or made cognitively compatible through human factor intervention – can significantly contribute to road safety. For instance, in the M8 long stretches of road often have very few or no road signs. However, the A71 has many road features in a short space, such as roundabouts, traffic lights, badgers signs, direction signs, and speed limit signs. Despite this, the richer SER nature of the M8 when compared to the A71 has been proven time and again.
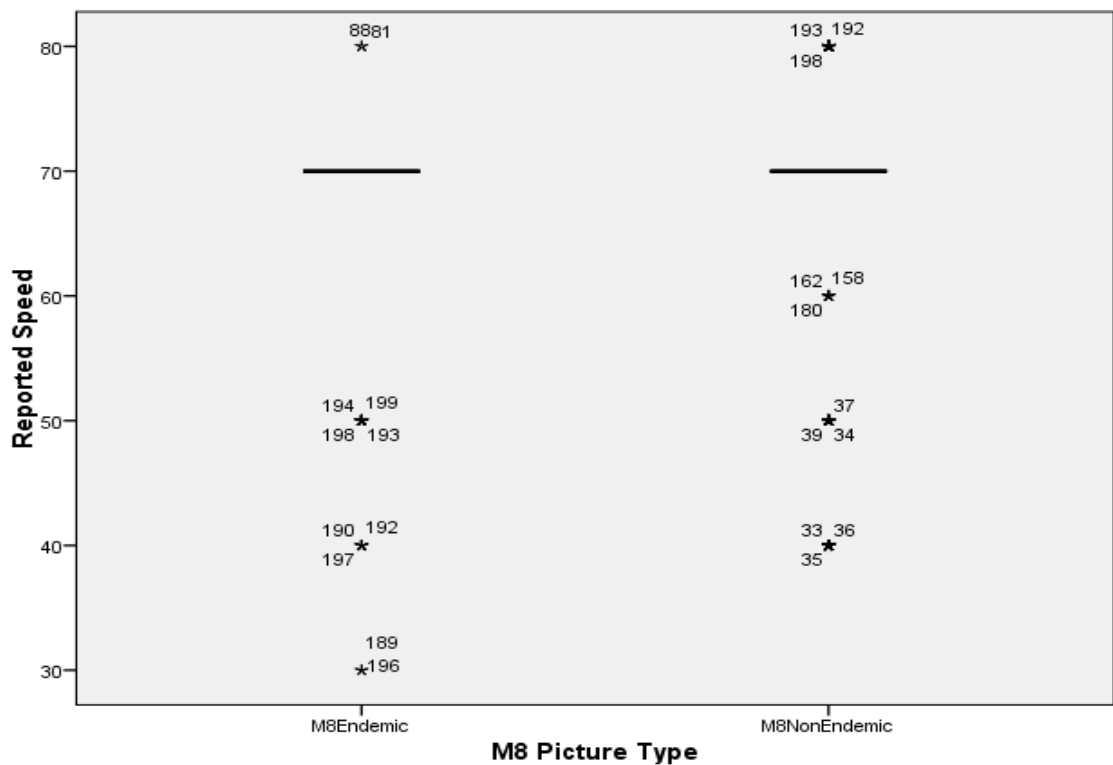
Figure 8.5. Box Plot of Difference between Absolute Posted Speeds and Reported Preferred Speed on the M8

Figure 8.5 can be understood in light of Table 8.1. The range of speed is higher in the M8 endemic category than in the M8 non-endemic category. All participants have chosen a speed of above the mean (M8 endemic adopted = 70 mph, mean = 67 mph; M8 non-endemic adopted = 70 mph, mean = 66.20). This again indicates the homogeneous nature of the motorway. The box plot of absolute speed on the M8 clearly shows a ceiling effect (shown in Figure 8.6). Ceiling effects occur when a test is relatively easy for participants, and a substantial proportion of participants obtain maximum or near-maximum scores (Uttl, 2005).
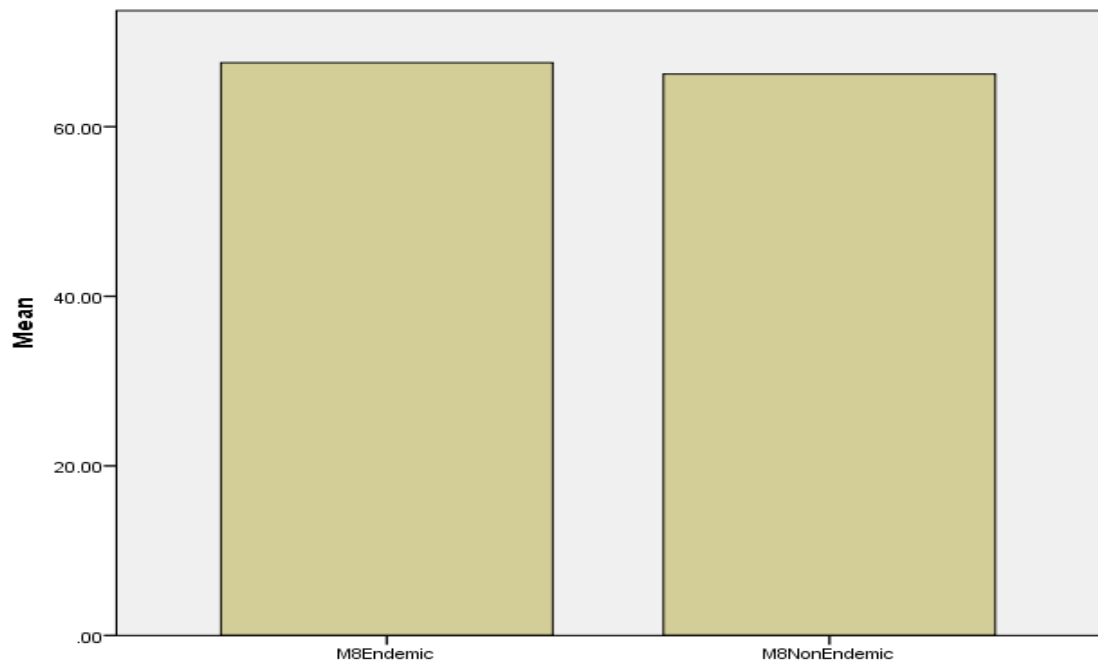
Figure 8.6 Bar Graph of Ceiling Effect of Preferred Speed on the M8

In this case, the posted speed is the maximum legal speed, such that subjects were reluctant to indicate speeds higher than 70 mph. This produces scores compressed in upper end of the performance. A test ceiling is the upper limit of the test. In the box plot shown in Figure 8.6 participants scored around the upper limit (70 mph) of the speeds provided. M8 is homogeneous within its road category and has a clearly defined purpose; in other words, it has high functionality. The design of the road is such that it evokes drivers to select a speed nearer to the speed limit. Like all tests suffering from ceiling effects, the box plot does not show a very clear effect of IV on DV.

An interesting observation which can be made from Table 8.1 is the increase in preferred absolute speed of the M8 endemic category (+1.35 mph). Thus, endemic features in the M8 condition are able to produce driver behaviour closer to speed limit, which may constitute an advantage as low speed is not always preferable. For example on the M8 (or any motorway for that matter) slower speed will have negative effects. Low speed is reported to be the cause of major hazard leading to rear-end accidents (Martens, 1997). In a motorway carrying moderate to heavy traffic, slower speed will lead to traffic congestion, restrict traffic flow and reduce road capacity (Martens, 1997). Without (or with less) endemic features the M8 can lose its functionality as drivers may

196

choose to drive at a lower speed by congesting traffic locally and increasing travel times.

The most striking results can be seen in the variance and standard deviation (in Table 8.1). Variance of the M8 non-endemic category is 130.211 (S.D= 11.41) and variance of the M8 endemic category is 61.187 (S.D= 7.86). Without (or with less) endemic features the variance of speed more than doubles in the M8 non-endemic category. Such high variance (than is found for the A71) can be attributed the ceiling effect of the M8. Since the range of speed in the M8 non-endemic category is lower than that for the M8 endemic category, this indicates that each participant interpreted the road condition differently. However, a larger range and less SD indicate more homogeneity.

The posted mean speed limit of the M8 endemic category is 68 mph and the mean adopted speed is 67.6 mph, such that drivers drove 1.6 mph below the speed limit. Similarly, the posted mean speed limit of the M8 non-endemic category is 70 mph whereas driver chose to drive on 66.55 mph, thus on average drivers drove 3.45 mph below the speed limit. This can be explained by several factors. First of all participants might have interpreted the road to be far more dangerous than it is in reality, influencing them to choose a lower speed. The second consideration is a statistical one. Since we are discussing mean speeds which are affected by extreme values, one or two participants might have scored extremely low and consequently decreased the average speed. Finally, despite being non-endemic, this category still forms a part of the M8 which can be considered overall as a SER. A SER 'elicits safe driving behaviour simply by its design' (Theeuwes & Godthelp, 1995). Therefore, participants chose to drive on a lower speed limit thereby maintaining a safety margin. This is completely opposite to findings for the A71, where participants chose to drive above the speed limit.

8.5.2 T-test for Absolute Speed on the A71 and M8

Table 8.2. Paired Differences for Absolute Speed

Paired Categories	Mean	Standard Deviation	Effect Size (r)	T-value	df	P value[1]
M8 endemic & M8 non-endemic	1.35	14.58	0.091	1.30	199	0.192
A71 endemic & A71 non-endemic	-2.12	13.38	0.156	-2.24	199	0.026
A71 endemic & M8 endemic	-19.42	12.11	0.849	-22.67	199	0.00
A71 endemic & M8 non-endemic	.18.07	12.90	0.814	-19.81	199	0.00
A71 non-endemic & M8 endemic	-17.30	10.69	0.799	-18.79	199	0.00
M8 non-endemic & A71 non-endemic	-15.95	15.14	0.788	-18.06	199	0.00

[1]*2-tailed significance (p<0.05)*

No significant difference of reported absolute speed is detected between the A71 endemic and A71 non-endemic conditions. Similarly, no significant difference is found between M8 endemic and M8 non-endemic conditions. Therefore, cognitive salient features did not elicit any change in absolute speeds. The difference in reported and absolute speed between the A71 and M8 are significant across all three conditions, as shown in Table 8.2. These differences confirmed the SER nature of the M8. This implies that the M8 is adequately designed to facilitate SA transactions between agents (markings, road surface, lanes, lights, signage, etc.) holding compatible SA.

8.5.3 Speed Variance
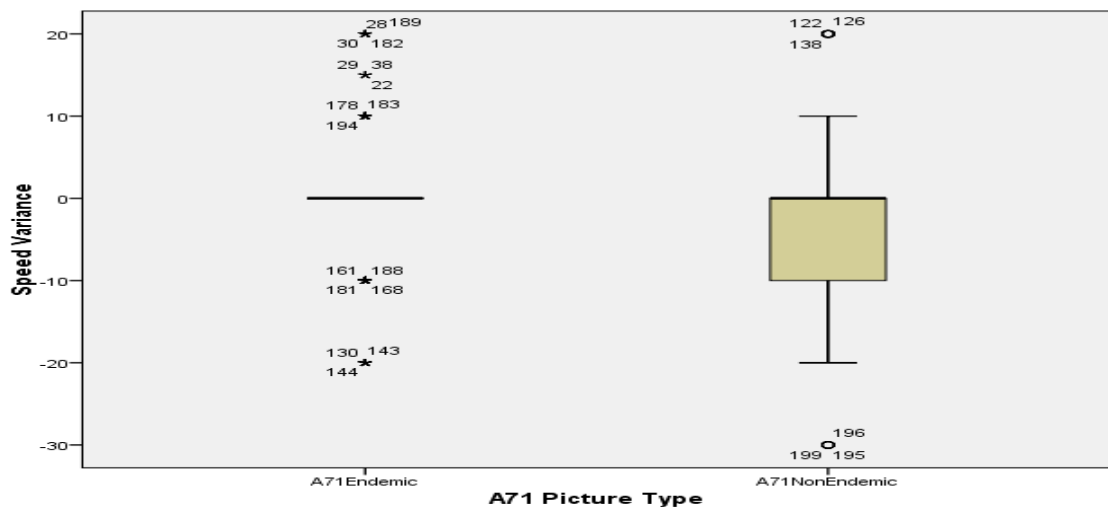
Figure 8.7 Box Plot of Variance for the A71

Figure 8.7 shows the box plot of variance for the A71. The A71 non-endemic category has a greater range of scores and much higher variance than the A71 endemic category. Higher variance means that the scores are further away from the mean and far apart from each other. This is exemplified by the box plot of A71 non-endemic category. The variance obtained by participants is very close to the variance of mean speeds in the A71 endemic category. Variance in these results demonstrates how closely aligned the present road design is with driver expectations. Higher variance will indicate greater discrepancy between what a car driver wants or expects, and what he gets in reality.

Table 8.3 Descriptive Statistics of Speed Variance

	A71		M8	
	Endemic	**Non-endemic**	**Endemic**	**Non-endemic**
N	200	200	200	200
Mean	-.7750	-3.2500	.4000	3.1500
Median	.0000	.0000	.0000	.0000
Standard Deviation	8.06845	9.97169	3.86194	13.83962
Variance	65.100	99.435	14.915	191.535
Range	40.00	50.00	50.00	140.00

Variance of each participant for every picture was obtained by subtracting posted speed from adopted speed. By this method, 200 variance values were obtained for 20 each participants for each of 10 pictures. Table 8.3 clearly shows participants' preferred or adopted speed is greater than the posted speed limit in the A71 when compared to the M8. This indicates a substantial mismatch of driver expectations, parallel to the previous results in this thesis. A self-explanatory road has greater consistency and homogeneity of speed at each level of road, and this is expected to have safety benefits for a given road environment (Charlton et al., 2010). This is demonstrated in the results. The sum of differences is higher in the A71 non-endemic category than in the A71 endemic category. The A71 non-endemic category has the highest sum of differences amongst all of the considered road conditions. Table 8.3 shows a constant decline in variance and standard deviation with rises in endemic features. Thus, the M8 endemic category has the lowest variance and SD, whereas the M8 non-endemic category has the highest variance and SD across all the road types. Drivers are adopting a more consistent speed when they are on SER. Similarly, for the A71 the non-endemic condition has a higher SD and variance than its endemic counterpart. This suggests that the endemic features are guiding the actual speed closer to the speed limit.

From this it is clear that endemic features are increasing road functionality, i.e. the purpose or the practicality of the road. Unlike motorways, A-roads are not clear in terms of purpose, as many times they are used as trunk roads, city roads and access roads. Multipurpose roads lead to confusion in the minds of drivers, incorrect driver expectations and inappropriate driving behaviour (Charlton et al., 2010). Thus, drivers often vary their speeds to a great degree which in turn inhibits the road being self-explanatory in nature. The results also point to an overlap of speeds within road categories, especially in the A71 non-endemic condition. This is again attributed to the poor SER characteristics of the A71. It is clear from Table 8.3 that the M8 endemic category has a lesser sum, SD and variance than the A71 endemic category.
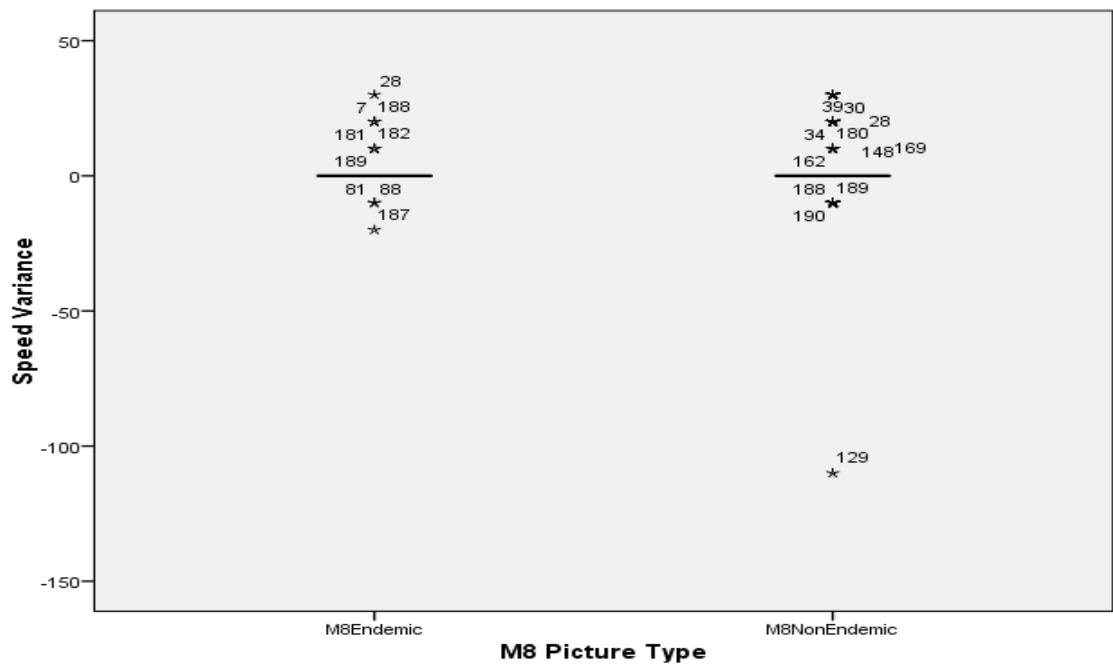
Figure 8.8 Box Plot of Variance for the M8

Like Figure 8.6, Figure 8.8 also shows a ceiling effect for participants relating to the M8. The number of outliers is greater in the M8 non-endemic category rather than the M8 endemic category, thus drivers may adopt more extreme behaviour in the non-endemic section of M8. Since the speed is higher in a motorway, the variance differences between endemic and non-endemic categories are over double. Furthermore, the range of variances is also high for the M8. This means that some drivers have chosen to be overly safety-conscious and maintain a wider gap from the posted speed limit. Speed selection over 80 mph on the M8 is rare, so the probability of drivers choosing a lower speed is much higher. As mentioned above, a larger range indicates a subjective design of the road, thus leading to varied driving behaviour. The variances of individual participants are very close to the mean variance of speeds in the M8 endemic category. This implies that the design of the M8 endemic condition is able to successfully activate the appropriate schema of the participant, and the resultant speed limit is appropriate.

8.5.4 Summary of Speed Variance
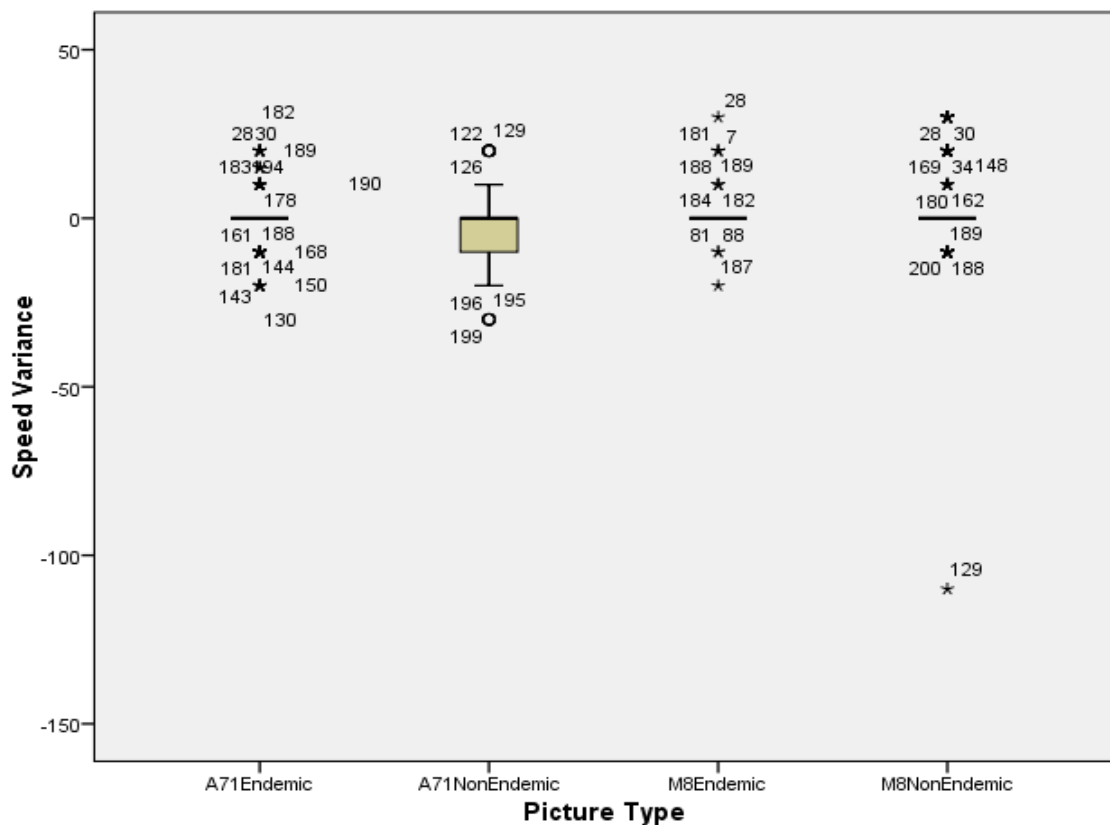
Figure 8.9. Box Plot of Variance for the A71 and the M8

Figure 8.9 and Table 8.3 show that the M8 endemic category is the most self-explanatory in nature. This section of the M8 has no fast transition from one road category to another (Weller & Detzer, 2010), unlike the A71. Furthermore, the M8 endemic category is able to reduce speed differences, and its transition to A-roads and end of motorway are marked very clearly, thus fulfilling the criteria of SER. Homogeneous behaviour leads to predictability (Materna et al., 2006). From the results, clearly the M8 endemic category is the most predictable in nature. Another feature of SER is enabling the driver to help correctly anticipate behaviour of other road users. This is better illustrated in the previous chapter covering verbal protocol analysis. For instance, the concept of 'roundabout' which occurs in the A71 and the M8 is connected differently for each of these roads. For the M8, 'roundabout' is connected to 'people', 'approaching' and 'lights'; however for the A71, 'roundabout' is connected to 'lane' and 'take'. Hence, the car driver on M8 receives plenty of signs as well as reinforcing

road signs to warn them of a change ahead. Driver on A71 are only informed of immediate actions to take. This in turn helps the driver to activate the appropriate schema and plan ahead.

8.5.5 T-test of Speed Variance

Table 8.4 Paired Sample T-test of Variances for the A71 and the M8

Paired Categories	Mean	Standard Deviation	Effect size (r)	T-value	df	P value[1]
M8 endemic & M8 non-endemic	2.47	12.12	0.200	2.88	199	0.00
A71 endemic & A71 non-endemic	-2.75	14.31	0.188	-2.71	199	0.00
A71 endemic & M8 endemic	-1.17	8.55	0.136	-1.94	199	0.54
A71 endemic & M8 non-endemic	-3.92	13.06	0.287	-4.24	199	0.00
A71 non-endemic & M8 endemic	-3.65	10.71	0.322	-4.81	199	0.00
M8 non-endemic & A71 non-endemic	-6.40	18.32	0.330	-4.94	199	0.00

[1]*2-tailed significance (p<0.05)*

Paired sample statistics (shown in Table 8.4) demonstrate that the increase and decrease in endemic features can lead to change in driver behaviour. More endemic features support the self-explaining nature of a road, and less variance in driver behaviour. The effect of endemic features on driver behaviour is strong in the M8, the A71, and overall. These results provide us with a core conclusion. Firstly, endemic features reduce absolute mean speed on the A71 by 2.13 mph. Secondly, endemic features reduce speed variance on the M8 (M8 endemic = 60.801; M8 non-endemic = 130.211). Variance is found to be strongly related to accident risk (Pisarski, 1986). On motorways, accident risk increases as vehicle speed deviates from the average speed. Garber and Gadiraju

(1989) state that the variance in speed depend on the actual driving speed. Motorways have higher fluctuation in speed than rural roads (Martens et al., 1997), although this can partly be explained by the difference in driving speed between various sorts of vehicles on motorways. For instance, lorries move more slowly than cars on average. If there are larger differences in speed between vehicles in a traffic stream, slower moving traffic may form an obstacle for faster traffic. For example, drivers are often found struggling to overtake the slower moving lorries on the M8.

8.6 Summary

This chapter validated the concept of cognitive salient features. This was done with a picture-rating task where cognitive salient features acted as IV and speed acted as DV. Speed was used as observable driver behaviour due to its strong association with road accidents (ETSC, 1995). The picture-rating task was used because it is cost-effective, requires minimal time, and has been successful with SER experiments in the past (e.g. Kaptein, 1992). Chapter 8 contributed toward fulfilment of the fifth objective and final objective of this thesis.

The results of the picture-rating task are discussed in two parts: reported speed and speed variance. CSF did not elicit any change in reported absolute speed although significant difference is found between the A71 and the M8 as a whole. CSF did show significant difference between the A71 endemic and A71 non-endemic categories, as well as the M8 endemic and M8 non-endemic categories. CSF are found to reduce speed variance. Speed variance is associated with increase in frequency of crashes (Virginia Department of Transportation, 2000). Higher speed variance also implies that there is high interaction with other vehicles thereby increasing the passing manoeuvers and opportunities for collisions (Virginia Department of Transportation, 2000). Designing below the posted speed limit will also lead to high speed variance (Virginia Department of Transportation, 2000). This is particularly seen for the A899 which is a dual carriageway with a 50 mph speed limit.

In summary, results regarding CSF are significant and when taken into consideration while designing can reduce speed variance of vehicles. The concepts underpinning these features are SER and SA. Cognitive salient features are actively involved in SA

transactions and hold compatible SA thus making a road rich in SA. Therefore road designs with fewer features but a high proportion of cognitive salient features can reduce road accidents. The succeeding chapter will use these cognitive salient features to develop a method to pinpoint SER hotspots.

Chapter 9 : Practical Tool for Transportation Engineering

9.1 Introduction

The thesis is written with the aim of laying the human factor foundation of user centred roads. This is done by using a well-established human factor construct SA; which is found to be critical to safety (Stanton et al., 2001). The substantial original contribution of this thesis is to prove that SA is at the heart of endemic features and deriving them through propositional networks increases its scientific robustness.

DSA purports that SA of a system is maintained through interaction between the system's different elements through SA transaction. Therefore in car driving road environment plays an important role towards maintaining SA of the road system. This has also been proved in Chapter 4 through SA requirement analysis and naturalistic study that car driver SA is made up of SA elements external to the drives relating to the road environment. A practical example of DSA in roads is SER.

The idea behind SER is that the road environment elements like signage or layout is able to communicate to its road users the "messages" required for safe and efficient driving (Evangelos et al.,2011). The term itself implies interaction between road roadside environment, trees, roundabouts and many more with road users (Evangelos Bekiaris & Gaitanidou, 2011). Driving is a combination of feedback and feed-forward mechanism. This means driving is a result of successful interaction which facilitates SA transaction through agents holding compatible SA. These agents are cognitive salient features as extracted in the chapters before. A successful interaction between roads, vehicle and driver will result in safe driving behaviour from the driver.

In order to choose a suitable model of SA to base this thesis on a comprehensive review of SA was conducted in Chapter 2. The review showed that DSA is the way forward to understand complex sociotechnical system and is extremely fitting for the aim of this thesis. The methodology of the thesis was devised through an analysis of SA methods in Chapter 3. It was concluded from chapter 3 that measurement has not kept up with the theoretical advances in SA models and definitions. It also showed that propositional

networks is the only methods which can explain the key components of DSA i.e. SA transaction and compatible SA among human and non-human elements of the road.

Chapter 4 discovered the information sources for driver and also put forward how it changed with different levels of driving. A secondary conclusion of Chapter 4 was that multimodal feedback is essential to safe driving and should be point of meaningful consideration for the transport authorities. Furthermore, chapter 4 also showed over an exploratory study that car drivers extract most information from the road environment. The source of driver expectancy comes from the road itself. Better SA transaction with the agents on road will mean stronger SER. A large scale controlled on road study was done (Chapter 5) and the results of Chapter 4 were confirmed. The inescapable conclusion from chapters 6, 7 and 8 is that cognitively salient features determine the SER nature of a road. Moreover, it also proves it is the cognitively salient features which illustrate the SA knowledge objects in a self-explaining road. A secondary conclusion is that M8 is more self-explaining in nature as it has adequately placed cognitively salient features than A71. Furthermore, the speed variance in M8 is lower in A71 suggesting a good correlation between driver expectation and M8 design. Thus again it is the road environment holds the key to safe and efficient driving.

The unavoidable fact of all the previous chapters is the success of SER is attributed to endemic features. These endemic features are indigenous characteristics of a road which enhances a road's functionality thereby increasing its predictability. However, all endemic features are not cognitive salient in nature (chapter 7). A combination of right number of endemic features and other important road characteristics make a road user friendly. Hence, cognitively salient features make a road far more self-explanatory in nature than just using endemic features itself. In other words, interleaving SA in the system can make a SER even more self-explaining in nature. This makes inevitable to define which road or road section in a network is more user friendly than others. Once known it will support the transport authorities to make intelligent investment and engineering decisions which can increase the safety, reliability and efficiency of a road. Hence, Chapter 9 is written to fulfil the last objective of the thesis which is to develop a tool to assess the cognitive compatibility of roads and validate it against real-world accident data.

9.2 Road Drivability

Drivability can be understood as 'drive' + 'ability'. That is, the ability to drive on a particular road. Determining drivability of a road will provide answers to questions such as how user-friendly a road is. How easy it is to drive on that road? To what extent the road conforms to driver expectations? In lieu of this *drivability is defined as the amount of SA in the road system. A higher drivability score will imply good SA transaction across all the agents in the road system. Consequently, a low score will imply potential of breakdown of SA thus leading to cognitive compatibility which can be viewed as SER hotspots. So these areas will be in need of immediate attention.*

SA centred SER enables the driver to adopt his/her behaviour in line with the current traffic condition without any additional effort. This form of road design is very objective so the driver is not left with a range of possibilities to evaluate and choose a suitable behaviour. If SER approach is not adopted and SA is not taken into consideration then uncertainty will exist. As long as uncertainties exist the driver cannot evaluate other possible decisions (Allen et al., 1971). Hence, in order to make rational decisions driver must eliminate all the uncertainties (Allen et al., 1971). This can happen if the road has SA embedded in it because it is a process which continuously gets updated as a drive drives the road. Therefore, driver expectation can be supported by SA centred SER. It is an "approach that intends to raise accurate expectancies in all the road users about how to behave, by designing and equipping roads in such a manner that they can be instantly interpreted" (Cocu et.al, 2011).

Drivability is dictated by design of the road and its infrastructure features. For example, narrow lanes make drivers drive on a lower speed. Also, roads enclosed with building and trees lead to reduced speed choices. In this regard the road drivability tool (RDT) will provides a measure of cognitive compatibility of a road or to what extent the road conforms to the driver's expectations or can influence driver behaviour. That is to say, how much feed-forward information the road can supply to the driver. Obviously, good feed-forward will imply the perceived image of the road is same as the objective state of the road. Currently, there are no desktop means to determine cognitive compatibility of the road. This poses as significant challenge to human factor researchers and the application of the theoretical findings into the real world (Salmon& Stanton, 2013). The

closest thing to measuring cognitive compatibility is assessing the road post SER intervention. This intervention only happens if an incident (in most cases many incidents) has taken place. At present drivability can only be assessed relatively late in the design lifecycle when it is difficult and most expensive to make any modification. Instead of waiting for something to happen and then gauge the SA inherent in the system, it is better to assess the drivability of the road before opening it to use or undertaking 'normal' interventions such as engineering and enforcement countermeasures.

9.3 Background and Application

The concept of drivability is not new to transportation domain. It already exists in rail transport. Hamilton et al., 2007 defined route drivability as the demand on the driver imposed by the infrastructure and environment of the road. Thus, drivability to a very large extent depends on environment design which in turn can be modified by self-explaining characteristics. Clearly, cognitive compatible self-explaining roads will have high drivability and consequently will improve driver SA, thus drivers will be adopting a safe driving behaviour automatically. Too many infrastructure features will increase the number of tasks to be performed by the driver in less time (Hamilton et al., 2007). This has also been supported by Charlton et al., 2010 who reported that too many endemic features have negative effect on driver. This makes it inevitable to determine those features which contribute towards driver SA and make a road system safe. For example, speed transition is found to be better controlled by rumble strips rather than road markings (Martens et al., 2007). Similarly, road markings in the centre of a road increase driving speed and reduction in road lead to decrease in speed (Martens et al., 1997). Such features are mined in the preceding chapters.

Some infrastructure features are more favourable and powerful in eliciting a particular behaviour than others so carefully placing these features on a road and integrating them into road design will increase the drivability of that road. This will also prevent from driver information overload. For example, multiple sign panels in one location (Lerner et al., 2003). SA centred road design i.e., integrating SA with SER approach can be a solution to this problem. It is likely that this is the only solution to mitigate driver information overload as traffic engineers are not unanimous to what constitutes driver

information overload (Lerner et al., 2003). Hence, engineering intervention cannot overcome this design problem. This thesis thus puts forward the cognitive features which can evoke favourable behaviour from drivers simply through design. RDT aids this in pointing out those zones which are inhibiting safe driving behaviour from car drivers.

9.4 Domain of Application

The broad approach is originally applied in railways to improve safety, capacity and efficiency of railway network (Hamilton, 2007). However, the method developed in this thesis is targeted at the route driveability of roads. In concept the tool is generic and could be applied to many areas where humans interact with technology, environments and infrastructures in any domain. For example, assessing usability of mobile phones, dashboard and any other area/ product which involves interaction with humans.

9.5 Procedure and Advice

RDT has been sequentially developed throughout the thesis. Chapter 1 identified that there are no quantitative methods to determine SA of a road, chapters 2 and 3 provided with DSA which forms the thrust of RDT and a multi method approach which led to extract CSF in chapters 6 and 7. In chapter 7 through social network analysis each concept was given a centrality value. A picture audit of A71 and M8 used in chapter 5 was conducted in chapter 8. This means screenshots of each section of the road were taken from google maps and on the basis of concepts in the picture centrality of each picture was calculated. An intermediate step here was to check the validity of CSF derived. This has been discussed in chapter 8 through the results of the picture rating task. Once validated, these pictures which are already quantified on the basis of centrality were matched with the accident data. Since CSF is found to reduce speed variability which is an important cause of accidents (Charlton et al., 2011) hence, sections of roads with CSF will have lower accident rates than their counterparts with no CSF.

RDT exists in two forms. One is doing it from the very beginning that is starting from route selection, data collection, extracting features, etc. Another version is a much

shorter version called RDT express. If using the RDT express version then steps 2, 3, 4, 5 and 6 (mentioned later in this chapter) can be skipped. The express version works with the cognitive salient features extracted from this thesis. The cognitive salient features have been validated twice in this thesis so it can be used for assessing drivability of similar roads.

Step 1: Identify the road to be assessed.

The first step in RDT is to carefully define the road under analysis. This can be a road with high accidents or traffic violations. This step actually depends on the transport engineers or relevant transport authorities who wish to investigate the cause of accidents/near misses on a particular road.

Step 2: Participant Selection

Since the study is under naturalistic paradigm so selecting everyday drivers is preferred. Thus from a population of everyday drivers an appropriate sample of drivers are selected. Caution must be exercised in in selecting a sample. After step 1 it is recommended to look up the government statistics of driver's age. The sample must mirror the high level summary of statistic trend of transport board. This step can be skipped if cognitive salient features are used from this research.

Step 3: Concurrent Verbal Protocol.

Once the road is selected participants should be clearly instructed what is expected of them (materials provided in appendix). Participants are required to drive and think aloud implying they have to speak aloud what are they thinking as they are driving. An example audio recording of verbal protocol is played to the participant. If possible a video recording should be shown to the participant as well. It is made explicitly clear to the participants that they should keep talking even if doesn't make much sense. The idea is to get the information processing behind driving task of a car driver. A trial run is done at this stage to check if subjects are able to do negotiate the route with no or minimal discomfort. This step can be skipped if an existing dataset is already available of a similar road.

Step 4: Data Collection

After step 2 and 3 and good trial drive performance the participants should begin to drive on the road selected in step 1. The whole scenario should be audio recorded by the analyst. It is recommended to record the vehicle diagnostics through a data logger. This will give a good picture of interaction between the three main agents of a road system: vehicle, driver and environment. If the experimenter accompanies the driver during the run then recording driver's overt behaviour will be incredibly useful. After completing a run the verbal protocol obtained is transcribe using notepad or any other text application.

Step 5: Structural Analysis

The verbal reports obtained provide insight to the driver's information processing when they are completing a task. Once transcribed the transcripts are uploaded to Leximancer™. "Leximancer™ is a data mining tool that can be used to analyse the content of collection of textual documents and to visually display the extracted information" (Leximancer™ Manual, 2005). In Leximancer™ a folder is made for every participant. In that folder transcripts obtained from the driver are uploaded individually. After this a concept map is generated for each participant. . The details of how Leximancer™ produces these networks are reported in chapter 4 and chapter 7. The semantic network also provides the number of times each concept is occurring. It is expected that these network will change in shape and content with different parts of the road. Step 5 will show that some networks are smaller and dense with knowledge objects close to each other while others are bigger with knowledge objects are far apart. It will also show that some knowledge objects are unique to a particular road or section of the road. The unique features are endemic to a road so they are called endemic features.

Step 6: Content Analysis

Step 6 will look into sociometric status, centrality and diameter of a network. That is how compact a network is and which knowledge objects are more popular than others. A longer diameter will imply that knowledge objects are loosely connected to each other and the interaction between them is poor. This is done in AGNA. AGNA is free social network analysis software which can be easily downloaded from internet. Every knowledge object is entered in AGNA along with its frequency as showed in

Leximancer™. Step 6 will reveal the content of a network in the form of centrality and density of a network. Many endemic features obtained from step 5 might have low centrality whereas many non-endemic features may have high centrality. Endemic features with low centrality will mean that despite being endemic to a road it is not important for driver SA. Step 6 will produce list of knowledge objects on the basis of centrality. These will be the cognitive salient features of a road.

Step 7: Validation Study

In order to test if the extracted features are at all cognitively salient in nature a validation study is required. This can be done two ways. First, is to conduct a route audit using images of the road scene. This can be done using google images corresponding to the road under analysis and placing the extracted knowledge in those pictures. Inter-rater reliability should be established at this stage between two independent raters to ensure sharing of themes which is this case are cognitive salient features. Every image is then scored on the basis of number of cognitive salient features on them as each feature has been assigned a centrality value in step 6. Finally, step 7 should produce a list of images varying in the number of cognitively salient features.

Another way to validate this is using retrospective verbal protocol. Concurrent verbal protocol can provide a very rich and interesting data source but sometimes are incomplete (Banks et al., 2014). It is possible during a concurrent verbal protocol participants may not report few elements but that doesn't mean that they are unaware of it (Bainbridge, 1999). This can be supplemented with retrospective verbal protocol. The route can be video recorded and played back to the same participants who participate initially for the naturalistic study. The participants then verbalise their thoughts afterwards on the basis of their task performance as seen from their video recording. After retrospective verbal protocol steps 5 and 6 are repeated.

Step 9: Prioritise Remedial Actions

Thus, results of step 8 will provide some 'SER hotspots' which need immediate attention of transport engineers to improve driver safety and make it self-explaining in nature. These hotspots are checked with the accident available. In this case it was checked with the accident of 2011 obtained from accidentdatabase.co.uk.

9.6 Advantages

1. As driving is an interaction of road environment, driver, vehicle and other infrastructure features, road drivability underpins this interaction. Moreover, it specifies where these interactions are strong and weak.

2. Road drivability has been developed from semantic networks and social network analysis. Each of them is well-established in the human factor literature and has been used widely.

3. Road drivability is easy to learn and can be used by non-experts.

4. Provides output in the form of SER hotspots. That is to say, the outputs of road drivability can be used to inform road design and interface design of driver assistance systems.

5. The networks which have contributed towards development of road drivability can be analysed mathematically. These networks underpin "activated knowledge" of a system. This information is immensely useful to transport planning authorities.

6. Outputs of previous studies can be reused to assess the drivability of a road. For example, cognitive salient features extracted from on-road studies in this thesis can be used again and if required can be reanalysed.

9.7 Disadvantages

1. Can be time and resource intensive to conduct a route driveability study from scratch, but the outputs from previous studies can be reused if the context is similar. Nevertheless, needs considerable preparatory work.

2. Road drivability does provide a very good basis for the practitioners to work with. However, links need to be established between identified SER hotspots and driver error. Also, links need to be established between SER hotspots and actual behaviours and errors

3. May or may not be cost effective to perform method depending on accident performance of road to be analysed.

4. Requires subject matter experts to check reliability of validation study.

9.8 Related Methods

Route Drivability by Ian Hamilton, 2007 was developed with the support of Network Rail to assess the drivability of a railway track. The infrastructure features such as signalling schemes, line speed, radio markers and many more define a train driver's task. More infrastructure features will increase the number of tasks to be performed within a limited time (Hamilton et al., 2007). This will increase the mental workload of the train drivers. So, Hamilton defined drivability of a railway track in terms of mental workload.

Driveability Index given by E. Bekiaris, Amditis, & Panou, 2003 defines drivability of a driver instead of the road. According to Bekiaris et al., 2003 car driving is a cognitively demanding task and depends on a range of physical, mental and psychological factors. These factors are not static and can significantly affect a person's driving performance (Bekiaris et al., 2003). The authors of drivability index are of the idea that drivability of a driver changes over time and hence should be assessed regularly the results of which will inform the licence issuing authorities. Driveability index depend is dependent on five factors individual resources, knowledge and skill of the driver and environmental workload and risk awareness (Bekiaris, et al., 2003). The latter two are deemed to be the most important factors of drivability. This index fails to illustrate to the fullest the effect of SER on a driver.

The drivability tools mentioned above do not take into account any detail of the driver-infrastructure interaction. Hamilton's route drivability recognizes the interactionist perspective of driving but doesn't appreciate this fact. Workload is the demand placed on the humans while negotiating a task. According to Hamilton's route drivability infrastructure features places on the train drivers. So, this is essentially a feedback phenomenon. However, as mentioned before driving is a mix of feedback and feed forward processes. So route drivability is actually giving a measure of feedback of the railway track to the train driver. There is no element of predictability in it which can inform authorities to make intelligent decisions before use. Feed forward element of driving can be supported by SA. SA is defined as "activated knowledge of specific task at a specific time" (Stanton et al., 2006, p.1291). In order to know which route is more drivable than other or how make a route drivable it is integral to know which knowledge is getting activated and how. That is, how the human and non-human elements of an

environment are interacting with each other in activating that 'knowledge'. This information can help to place the route elements intelligently to support drivers.

There are other accessory methods which are also related to road drivability and contribute towards its development. Task analysis and helps to decompose the tasks to the very basic level and helps to understand the complexity of the tasks. Requirement analysis helps to determine the driver's needs and source of those needs at every task level. Verbal protocol analysis is also used as an accessory method to collect driver's driving protocols which are analysed by Leximancer™ and AGNA. As control measures driving style questionnaire, mental workload, and SA rating technique can be used.

9.9 Approximate Training and Application Times

The analysts involved preferably should have some basic understanding of SA, psychometrics and schema theory but not mandatory. After data collection the transcription of verbal transcripts can take significant amount of time. Time of transcription depends on the sample size. Bigger sample will mean higher amount of transcription time. In this PhD over 38 hours was spend for transcribing for 35 verbal transcripts. Training time should not exceed 2 hours. In this thesis following the extraction of endemic features and validating the method, it took 6 hours to assess the drivability of the route. The application time can be high if the route to be analysed is long stretch of road. Caution should be exercised in putting the correct number of endemic features in the correct picture. Since many sections of the road look identical it is very easy to get confused with the number of endemic features.

9.10 Reliability and Validity

Reliability refers to the consistency of a measure. That is if the test produces same results repeatedly. Satisfactory inter rater reliability is established between two raters for the evaluating the drivability of A71 and M8. For IRR 1, rho = 0.841 and for IRR 2 rho = 0.868. Both the values significantly correlated ($p > 0.05$ for IRR1 and IRR2 with n=20) with each other. Future studies should develop test-retest and intra-rater reliability. The test is found to be sensitive to changes in the number of endemic features. Future

studies will establish the test retest reliability of RDT. That is, to determine if RDT can produce similar results relatedly and consistently. Although more testing is needed the results so far are encouraging

Validity refers to whether or not the test measures what it claims to measure. Reliability can be understood as precision and validity as accuracy. Construct validity can be defined as the correspondence between a construct which is unobservable and at a conceptual level with its purported measure which is at an operational level (Peter, 1981). RDT is established from semantic networks which has a long standing pedigree and has been used widely. RDT maps to robust SA and SER concepts.

Ecological validity means how well the study can be related or reflects real life. Chapter 8 shows how speed choices differ with number of endemic features. Furthermore naturalistic studies have contributed to the construction of propositional networks thus again increasing the ecological validity of RDT. Predictive validity is the extent to which a score on a scale or test predicts scores on some measure (Cronbach & Meehl, 1955). RDT results can be correlated with the accident data. SER hotspots are those areas having high accidents.

9.11 Tools Needed

Road drivability can be conducted using pen and paper at a very simple level. However, audio and video recording devices can be used during data collection phase which is typically through verbal protocol. The verbal protocol is transcribed using standard office PC tools. The detailed analysis of the protocols is done with Leximancer™ and AGNA. The images of the road under analysis are also required. Microsoft excel is needed for constructing heat maps but this can also be done in Microsoft paint.

9.12 Flowchart
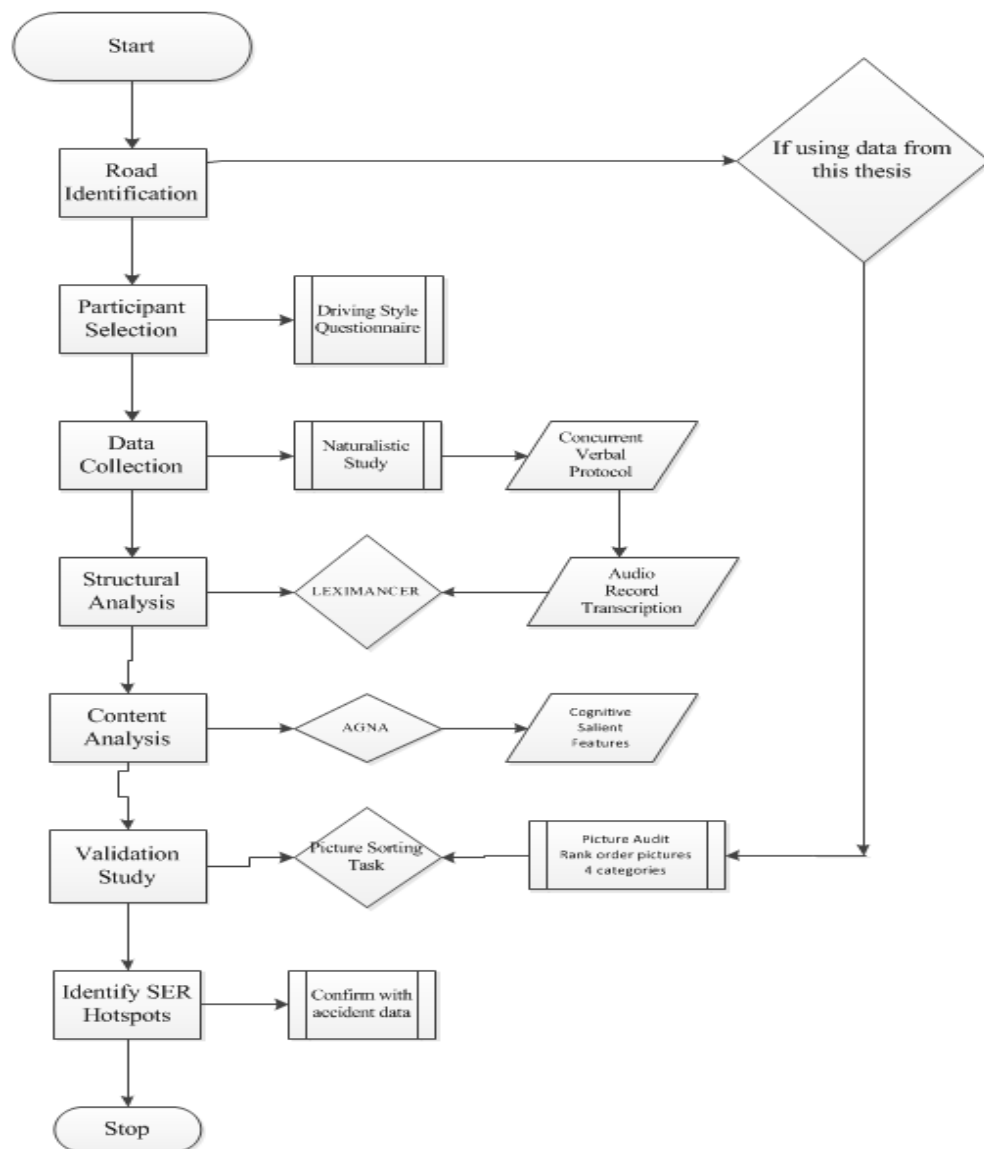
Figure 9.1 RDT process

9.13 Example

The following example illustrates the steps performed above from the start. The data collected or obtained has been used and analysed throughout the thesis.

Step 1: Identify the road to be assessed.

Details of the route are in Chapter 5. A 22.5 mile road was selected for this study. 11.9 mile long A71 served as an experimental condition. 10.6 Mile long M8 served as a

control condition and 1.3 mile long internal Heriot-Watt Road served as practice route. A71 started from Calder Road had two major roundabouts; Oakbank and Lizzie Brice. There are 3 speed limits on A71. It starts from 40 mph, and then goes to 30 mph as it goes through Wilkieston village and finally goes to 50 mph. After Lizzie Brice roundabout A71 turns to a dual carriageway A899 with speed limit of 50mph. M8 is a standard UK motorway with 4 lanes, 70 mph speed limit and 3.45 wide lanes. This route was chosen because of its accessibility as most of the participants identified were employees of Heriot Watt so it was convenient for them to participate. The selected route has variable and unconventional speed limits and the road network frequently report of accidents. So, this route was chosen to measure its extent of safety.

Step 2: Participant Selection

A large scale participant recruitment program was carried out at Heriot Watt University to recruit participants. 36 people responded with interest but 35 verbal transcripts were taken into account. One was discarded because the vehicle used was a people carrier. Interested participants were initially contacted by phone and then sent an ethical approval form, driving style questionnaire and driver declaration form.

Step 3: Concurrent verbal protocol

On meeting the participant at Heriot-Watt reception the completed forms were collected, their insurance status and MOT were checked. Drivers were instructed to speak while driving and keep speaking even if doesn't make much sense. A sample recording from an official who was a police driver at some point was played to participants as an example. A 1.4 mile trial run was conducted inside Heriot-Watt university campus. All the drivers were compensated for their time in the form of £20.00.

Step 4: Data Collection

Data was collected in the form verbal protocol which was transcribed to form verbal transcripts. After finishing the drive on A71, a short break was taken where participants completed NASA TLX and SART. These were filled again by the participants at the end of the study. So finally, 4 datasets were obtained for each of 35 drivers. They are NASA TLX, SART, verbal protocol and DSQ.

Step 5: Structural Analysis

The verbal transcripts were uploaded to Leximancer™ which generated semantic network for each participant. So, 70 semantic networks were obtained for 35 participants. Each participant had one network for A71 and one for M8. A list of unique objects of A71 (table 9.2) and M8 (table 9.1) were noted from this stage. These unique knowledge objects were endemic features. As expected from Chapter 4 pilot study M8 had higher percentage of endemic features.

Table 9.1 Endemic Features on M8

Concepts	Occurrence	Percentage
outside	289	97%
motorway	104	35%
overtaking	72	24%
past	68	23%
doing	61	20%
inside	61	20%
lorry	55	18%
pull	23	18%
bridge	53	18%
driving	50	17%
fast	44	15%
time	39	13%
mile	33	11%
bypass	32	11%

Table 9.2 Endemic Features on A71

Concepts	Occurrence	Percentage
limit	198	37%
gear	124	23%
clear	100	19%
carriageway	93	17%
straight	92	17%
camera	81	15%
turning	75	14%
slowing	71	13%
bend	67	12%
area	62	12%
markings	59	11%
vehicles	48	9%
take	52	10%
vehicle	48	9%
middle	43	8%
trees	42	8%
roads	57	11%

Step 6: Content Analysis

It was noticed from step 5 that some knowledge objects despite being non-endemic had high frequency rate. Thus, a social network analysis using AGNA was conducted to determine which knowledge objects more central than others were contributing towards the road's SA. Figure 9.2 shows the analysis. The end of this step provided cognitive salient features which are a mix of some endemic features that is unique knowledge objects and other important concepts.
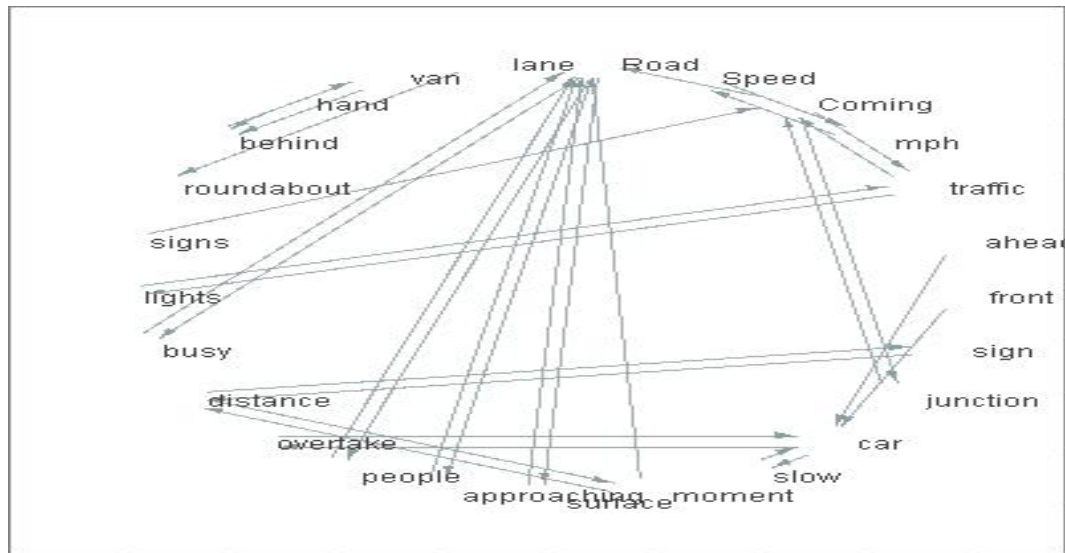
Figure 9.2 AGNA output of A71

Step7: Validation study

Picture Sorting Task was performed as a validation study because of its success with SER studies. Google images of the chosen route of A71 and M8 were collected and rank ordered on the basis of number cognitive salient features. These provided 4 lists of 10 pictures each. List 1: A71 high cognitive compatibility, List 2: A71 low cognitive compatibility, List 3: M8 high cognitive compatibility and List 4: M8 low cognitive compatibility.

Step7: Participant Re-recruitment

British drivers with full driving licence voluntarily participated in this study. 20 participants participated. All the participants were asked to summon in a room and the above mentioned 4 lists of pictures were shown to them for 5 seconds. Subjects were required to indicate their preferred speed choice within 5 seconds. Thus 20 subjects reported their preferred speed choice for 40 images. Stage 7 showed the difference between preferred speed and posted speed for all 4 conditions. It also showed the variance of speed in A71 and M8.

In case of using the existing endemic features, then this is the step to validate the endemic features mentioned in this thesis to the road under analysis. If they don't validate, i.e. show poor IRR then it is recommended to start the data collection for the route under question and then apply RDT.

Step 8: Prioritise Remedial Actions.

From the results obtained so far a heat map so to say is constructed of the route taken. Sections of road where the difference between the posted speeds are higher are regarded as less cognitive compatible or less user friendly and as areas where the SA transactions are poor. These sections are marked as red. Green represent road which are fairly cognitively compatible and difference between adopted and preferred speed is low. In Figure 1, the drivability score (which is summation of centrality) is 134.792 however in Figure 2 the drivability score is 200.890. Step 8 revealed the drivability of the 22.5 mile long route selected.

Figure 9.3 Road section marked red in the map

Figure 9.4 Road section marked green in map
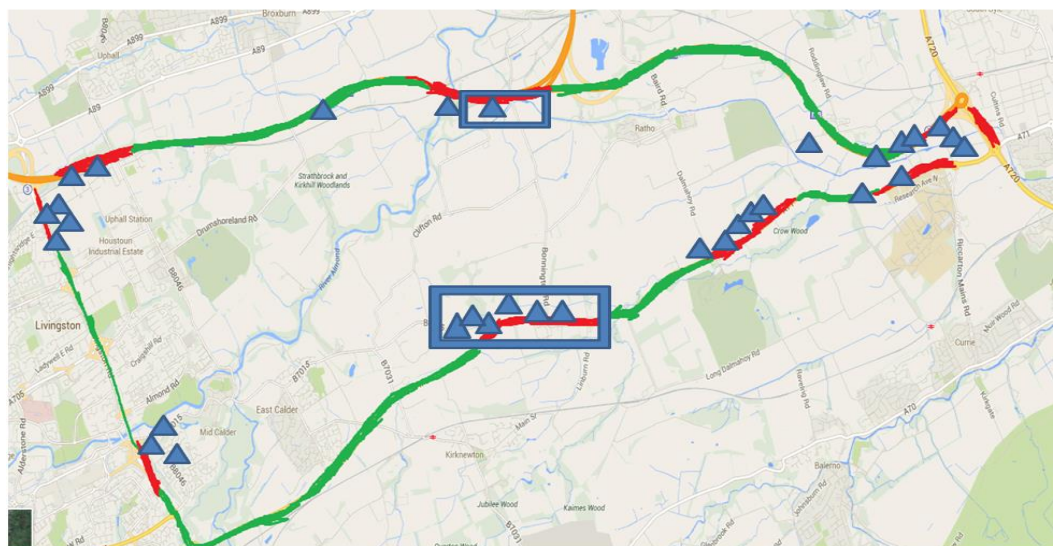
Figure 9.5 Cognitive Map of Experimental Route

Figure 9.5 relates to the accident data of 2011 and risk rating of M8 and A71 obtained from EuroRap based on accident record of 2006-2010. Figure 9.5 essentially displays a cause and effect relationship. The blue triangles show accidents which occurred in A71. The blue squares represent the medium risk levels associated with the road on the basis of number of accidents. The red areas obviously show zone of cognit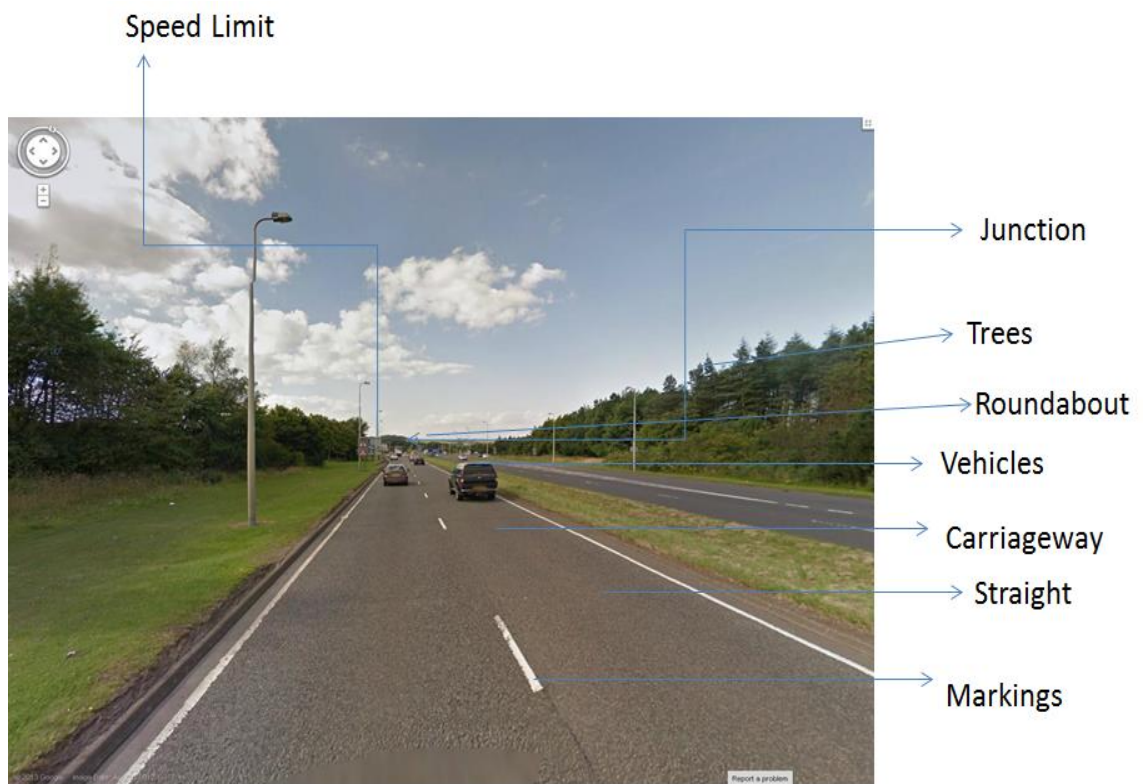ive incompatibility. So, the squared part of A71 i.e. the 4.5 mile road from Dalmahoy Road until the beginning Overton Crescent has had 14 accidents from 14 January 2011 to 23 December 2011 out of which 5 were deemed to be serious. This can be attributed due to poor SA transaction between agents of a system leading to cognitive incompatibility among drivers and the road.

The validity of RDT can be further understood from accident data of 2011. The route starts with A71 marked red which means cognitively incompatible. Three slight accidents are reported in that area from March 9 to July 25, 2011. The accidents involve 2 young males, a young female and 75 year old male. The speed limit in that area is 40 and there on progressing to 50 mph and the road is a single carriageway. 4 slight accidents were also reported in A899 which is a dual carriageway. 6 Accidents are reported from the selected section of M8. The end of M8 and beginning of A720 has 6 more accidents with one being serious. First point to make is a methodological one. Since the marked areas of low cognitive compatibility reflect the 2011 accident records it contributes towards high predictive validity of RDT. Second point to make is on the location of these accidents. A899 has a speed limit of 50 mph which is not a standard speed limit for other dual carriageways. So, A899 is heterogeneous across other dual carriageways in terms of speed but homogeneous in terms of layout. This results in conflicted schemas and mismatched expectancies for the driver. Thus, in case potentially demanding situation such as approaching the Livingston East Roundabout where SA transactions are intense accidents are higher.

The accident records of 2011 show that the 6 accidents in M8 are clustered around the end of M8. M8 ends to start the 6 lane Edinburgh City Bypass with one of them being fatal at rush hour (1753 hours). This transition will mean drivers will require higher interaction with the infrastructure features. In this situation confusing signage will increase cognitive compatibility. It will also mean in rush hour due to high volume of traffic there will be a rapid interchange of genotype and phenotype schema in the

drivers. Additionally, absence of hard shoulder and crash barrier will make the road unable to reduce the severity of accidents if any.

These results can identify the danger zones and accident black spots which are of immense use not only to transport engineers but also for insurance companies. The later can provide a quote on the basis of road they often commute on. If the road is not drivable than then can be asked for a high end policy.

Chapter 10 : Conclusions and Recommendations

10.1 Introduction

The overall aim of this research is to lay the groundwork for user centered road design through the adoption of a combined Distributed Situation Awareness (DSA) and Self-Explaining Roads (SER) perspective. This was achieved using a combination of desktop, naturalistic and classroom studies. A summary of main conclusions derived from the overall body of research is presented below, along with its wider implications.

10.2 What was done?

The research began with a gap in knowledge. A body of literature has been growing in the Self-Explaining Roads (SER) domain from leading authors such as Green, (2009), Charlton et al., (2010), Weller et al., (2008) and, Mackie et al., (2013). SER is based on the expedient that design should evoke "*correct expectations and driving behaviour from road users to create a safe and user friendly road network*" (Mackie et a.,2013, p.742). Despite some very apparent overlaps and a need for a more robust theoretical underpinning, the concept of Situational Awareness had yet to be applied to SER. This fusing of two disparate strands of research is a key innovation and a valuable way to extend the approach.

The twin concepts of SER and SA were explored, where it became apparent that a theory of SA had to meet a number of criteria for it to be suitable for use in an SER context. These were as follows:

- Acknowledging the fact that roads work as a system within which human and non-human agents interact.
- SA is an emergent property of these interactions meaning that it is a process and not a product.
- To be able to assess individual agent SA from within a multi agent system requires knowledge of the interactions between agents in the system.

227

The model best fitted to these criteria is a recent development called Distributed Situational Awareness (DSA). DSA is defined as "*activated knowledge of a specific task at a specific time within a system*" (Stanton et al., 2006, p.1291). It is novel because it views SA at a system's level and views it as an emergent property of a system involving human and non-human artefacts; a sociotechnical system. DSA, therefore, is referred to as the collective awareness of the system in which the agents are working.

Coupling DSA to SER created important new opportunities to explore the systematic relationship between drivers and the road environment. These became manifest in the course of a pilot study which was published in 2013 in the peer-reviewed journal Safety Science (and reported in this Thesis in Chapter 4 and elsewhere). In it an approach to enable endemic features of a road to be extracted using propositional networks was developed. The work formed the basis of a much larger Naturalistic Driving Study (described in Chapters 5, 6 and 7). In it a large pool of drivers who matched the demographic profile of Scottish drivers was recruited. The study required them to drive on a real-world test route comprised of roads local to Heriot-Watt University's Edinburgh campus (the A71 and M8) and a wide range of measures was extracted. Using the DSA approach the key innovation was being able to extract a number of cognitive salient features, those elements of the built environment which are instrumental to a road being self-explaining and for it to 'afford' correct speed behaviour.

Do those cognitively salient features really have an effect? In the validation study reported in Chapter 8 it was possible to demonstrate that, yes, they did work. A group of 20 participants undertook a picture rating task (a well-used paradigm in SER research) and the results showed that roads which contained more cognitively salient features were associated with reductions in overall speed and/or reduced speed variance (depending on whether the road was a motorway way or A-road). The practical implications of these findings are made manifest in a practical method that engineers can use to quantitatively and reliably assess a road for its drivability.

10.3 What was achieved?

Chapter 1 set out five core objectives for the program of research described in this thesis to achieve. Were these objectives met?

Objective 1: Explore different aspects of situation awareness and select a model most suited to the systemic properties of road infrastructure.

This objective is achieved in Chapter 2 through a comprehensive literature review. Chapter 2 provided a definition and model of SA which is followed throughout the thesis. Endsley and Endsley-type models have tended to dominate the SA scene by adopting a 'SA in the head' approach. This trend is also seen in the measurement aspect of SA. This thesis puts forward an argument that this model is not best suited to the specific research questions around SER. This is demonstrated to some extent by a mixed method approach of VPA, SART and naturalistic studies, all of which contribute to the achievement of the remaining 4 Objectives.

Objective 2: Determine ideal source of driver SA and compare how driver's information needs match the results gained from real world data collection.

Chapter 4 describes how Objective 2 was met. Driver's information needs were determined through an SA requirement analysis of 1100 driving tasks. The result indicated that drivers need multimodal feedback instead of visual feedback alone. The results also showed that SA is formed of elements in the road environment. This was further confirmed from the exploratory naturalistic study. The results of Chapter 4 demonstrated the fittingness of DSA in road transportation.

Objective 3: Understand the self-explaining characteristic of different British roads and reveal the link between SER and SA.

This objective is achieved in the second part of Chapter 4. The exploratory study showed that motorways are the most self-explaining in nature and that the SER nature of the road decreases as we go down the road hierarchy. As mentioned above, Chapter

4 was successful in extracting the endemic features through a DSA approach, therefore, establishing a robust link between SER and SA.

Objective 4: Extract cognitively salient 'endemic features' required by SER approaches using distributed situation awareness methods and validate them.

The results from Chapter 4 were encouraging enough to scale up and attempt to achieve objective 4. This is described in Chapters 5, 6 and 7. A naturalistic study was conducted from scratch. This meant selecting a suitable sample size, choosing a representative sample, obtaining ethical approval and finally recruiting participants. Thirty five subjects took part and each provided a verbal transcript as they drove over the A71 and M8 test route. The DSA analysis of these transcripts revealed the presence of cognitively salient features which are pivotal to ensuring cognitive compatibility and maintaining road safety.

Objective 5: Develop a tool to assess the cognitive compatibility of roads and validate it against real-world accident data.

The final stage of the thesis was to test the validity of the results obtained from the naturalistic driving study. For this, a smaller validation study with 20 participants took place to check the effectiveness of the cognitive salient features extracted from the naturalistic driving study. The thesis then integrates all these results into a road drivability tool (RDT) which can specify areas high and low on cognitively salient features. This enables areas of cognitive compatibility and incompatibility to be identified. This step not only contributes to the body of knowledge but provides engineer's with a user-centered view of the built environment.

10.4 What was found?

The research establishes five crucial results which are detailed below.

10.4.1 Distributed Situation Awareness is the Way Forward

After a careful review of situation awareness literature it was found that SA is a contentious concept. There is no universal definition or model of SA since its inception

in World War 2. From a plethora of models DSA was chosen for the thesis. The foundation of DSA was laid by Stanton et al., (2006) which was later developed into a full-fledged model by Salmon, (2008). According to DSA, situation awareness is an interactionist process which emerges out of a successful SA transaction among various agents holding compatible SA. DSA is founded on distributed cognition, schemas and the perceptual cycle of cognition. There are several reasons which have made DSA a clear choice for explaining road transportation SA. These are detailed in Chapter 2. The main reason DSA has been chosen over individual and team SA models is because DSA views SA as spread around the environment. Driver, technological devices, vehicles, artefacts and any other component of the road environment embody SA to some degree. It is not limited to the SA of drivers only. The DSA model has two important characteristics; compatible SA and transactive SA. DSA depends on the SA transaction between agents and not the amount of SA held by the individuals. On a road, SA transaction is possible even if one agent has negligible SA as the other agents can compensate for it.

A mixed method approach was adopted to attain the aim of the thesis. A method review revealed that SA measurement is a difficult task and the psychometrics of SA tests are hard to establish. Furthermore, studies on SA psychometrics are scarce which make it all the more challenging. The methods review also pointed out that the majority of SA methods are developed from individual SA models and thus are unsuitable for the current problem. Naturalistic studies following an on-road paradigm are used for deriving the key results, and DSA works well in these contexts. Naturalistic studies enable the researchers to unobtrusively record data during every day driving – using DSA methods means such studies do not need artificial pauses or environments, therefore offering complete ecological validity (Walker et al., 2002, Walker et al., 2009, Young et al., 2013)

10.4.2 Link between Self-Explaining Roads and Situation Awareness

SER makes use of number of principles, such as affordances, prediction and schemas (Walker et al., 2013). SA can unify all these concepts (Walker et al., 2013) and can contribute to the scientific robustness of SER. SA requirement analysis results also supports the relationship between SER and SA. The results of the requirement analysis

establish the primacy of the road environment and why, by extension, a road is self-explaining to the driver. Propositional networks specify which infrastructure characteristic make a road easily interpretable to its user. The structure of the network describe the knowledge underpinning driver SA, important concepts undergoing high SA transactions. The unique concepts obtained from propositional networks are the endemic features of a road. Endemic features are the cornerstones of creating SER (Charlton et al., 2010). Roads with endemic features have been successful in reducing speed in specific areas (Charlton et al., 2010).

An exploratory naturalistic study proved the link between SA and SER. The results confirmed that driver SA consists of design elements of a road, moreover, it was established that the SER nature of a road increases as we go up the road hierarchy. So the M8 (a motorway) is the most self-explaining and rural roads are least. This is mirrored in the SA network structure and content, which differs for each road type. For example, the M8 motorway has a high number of concepts concerning the road; A/B road with speed; urban roads with other traffic; and rural roads with road geometry. The exploratory study also demonstrated that SER roads are in agreement with driver expectations. For example, adopted speed on A/B roads were found to be 32.8% different from the posted speed limit whereas the difference between adopted and posted speed on motorways is 1.1% (Walker et al., 2013).

10.4.3 Cognitive Salient Features are the Key to Road Safety

Activating incorrect schema is the biggest cause of human error (Norman, 1981) and human error is the biggest cause of road accidents (Treat et al., 1977). Cognitively salient features act as triggers to activating a schema. These are, in essence, what endemic features are. Endemic features make a road what it is. Cognitively salient features are those characteristics of a road which, via active SA transactions and compatible SA, form the perceived image of the road to the driver. Determining these cognitive salient features is pivotal to road safety and any change in the road design should be done whilst keeping them in mind.

The notion of cognitive salient features is a novel one which is put forward in this thesis. It further adds to the association between SER and SA. Cognitively salient

features can be obtained by subjecting the verbal transcripts to a text analysis tool called Leximancer™, spotting unique and common concepts, or knowledge objects, across road types, and then sieving out the important ones from the rest. In this thesis the separation of important from non-important concepts is done through AGNA which is social network analysis tool. The concepts high on centrality are regarded as important and cognitively salient.

10.4.4 Is your Road Drivable?

The Road Drivability Tool (RDT) proposed in this thesis fills enables engineers to determine the extent to which a road is drivable or not on the basis of cognitive salient features on that road. A RDT score would determine if a road is cognitively compatible and has high levels of inherent SA. In essence, it will provide a numerical value of 'self explainingness'. RDT also confirms that more information is not always better. The method is theoretically robust. It built on the principles of distributed situation awareness and self-explaining roads and has foundations built on propositional networks and social networks. As mentioned below, the method is still in its infancy and needs further research in terms of refinement but it has fulfilled its purpose here. A RDT 'express' tool provides a shorter version of determining the drivability of a road.

10.5 Limitations

The aim of this research has been to determine the cognitive compatibility of road using the principles of SA and SER. The idea of combining the two might not be well accepted by SER and SA researchers at the present time, and there are inevitable questions.

The data collection phase relies on a think-aloud technique. Think-aloud is used widely among usability professionals and researchers across the world (Kumar et al., 2013) but it is not without flaws. For example, some thoughts are difficult to translate verbally (Shuck & Leahy, 1996) and mental processes are reported to be faster than verbalization speed (Bainbridge, 1990), so this leaves a gap in the verbalization of thought processes (Kumar et al., 2013). The result of this thesis has also been derived from English speaking participants. This is due to the fact that Leximancer™ only accepts transcripts in English but of course there are a considerable number of people who drive on

Scottish roads (not to mention roads the world over) whose first language is not English. This might limit the internationalization of the research, and perhaps one of the reasons DSA research has been limited to mainly English speaking nations, although this is changing. A solution to this is to translate every transcript into English and then upload it to the text analysis tool. This will not only increase the cost of the research but also make it far more laborious and time consuming. Something may also be lost in the translation process. This is a wider question for content analysis methodology.

The research also doesn't foreground the vehicle–road relationship. It focusses on the driver's view of the road environment and the effect of road design on driver behaviour. Of course, driving involves an input from the driver to the vehicle which initiates the electrical, mechanistic components of vehicle that translates into observable driver behaviour such speed and trajectory (Walker et al., 2006). The vehicle's resultant interaction with the road places stresses on its mechanistic and electrical components that influences driver behaviour (Walker et al., 2006). Hence it is fair to say than an implicit relationship exists between the driver-vehicle and road. The absence of vehicle behaviour data through a data logger prevents us from exploring the mechanistic properties of vehicles in cognitively incompatible areas of a road network, and the role of vehicles in making driver choose a speed on particular road. This is an interesting area for extending the research further.

10.6 Future Research

Throughout the course of this research a number of key areas of future research with SER-SA have emerged. A summary of these are given below:

10.6.1 Amalgamation of eye tracking and verbal protocol

Although verbal reports provide a very interesting insight into driver mental processes when they are completing a task they are not complete (Banks et al., 2013). Bainbridge (1999) argues that just because a driver does not report something while they are driving does not mean that they are unaware of it. Concurrent verbal protocols suffer from the criticism of implicit knowledge and how that may or may not emerge via the technique. Retrospective verbal protocol enables researchers to directly explore

subject's knowledge of a task (Bainbridge, 1999). Thinking aloud post task is an option which can be used to in conjunction with concurrent verbal protocol.

For future research a useful avenue to explore is eye tracking. Mapping eye tracking data with verbal protocols will undoubtedly aid in the design of SER. Eye tracking data can provide driver's field of view which differ significantly in a familiar and non-familiar road ("Attitudes, iniciativa social de Audi", n.d). Hence, this will enable the traffic engineers and ergonomists to appropriately place cognitive salient features. Eye tracking data in previous research shows that drivers don't remember everything they see, which means they may have difficulty in reporting it verbally.

10.6.2 Application and Development of RDT

A drivability tool has been presented in this thesis following a DSA design process. The next logical step will be to develop software support for the drivability tool. This obviously requires much more research around pinpointing a numerical value of a road which is indicative of a high drivability. Similarly, a standardized scale needs to be developed to supplement the RDT which can instantly provide the measure of drivability of a road. Like a spectrum. The software should ask the users to enter the observable road features, number of junctions, intersections, roundabouts, time of travel and speed limit. Future research should also validate the extracted cognitive salient features in this research so that they can be directly used by researchers without the time and cost of re-extracting them.

Once CSF is standardized many interesting RDT application can be undertaken. One of them can be to conduct the drivability assessment of a long stretch of road or whole road networks. This means existing roads can be checked for drivability scores which can provide transport engineers SER hotspots to work with. Further research can also be taken to arrange the road on the basis of a drivability hierarchy. Cognitive categorization of road and EuroRap risk rating for road should be consulted in conducting large scale drivability studies.

RDT also gives an opportunity to analytically prototype a design before constructing a physical road. For example, micro-simulation of an analytical design can be undertaken and checked for drivability for different sections of the road.

Human Factors as such is a very fruitful area where engineers and behavioural scientists work together. This can be further enhanced through further development of RDT. RDT works with CSF. In this thesis for example, drivability of a road to large extent is dependent on the CSF extracted. But CSF is extracted through content analysis. So, transport engineers need to exercise caution in interpreting and using these results and HF professional are needed to make sense of the complete meaning. CSF are those features which are pivotal to driver SA. They can also be those which are hindering driver SA. For example, "roundabout", "middle", "take" are CSF for A71 to drivers. This means drivers took the middle lane from the roundabout and they derived this information from the signage and the lane. Therefore the roundabout fulfilled its functionality by keeping the traffic flowing. The job for transport engineers here is to maintain this compatibility between drivers and road in future designs and roadwork's. Similarly, "approaching" and "marking" are CSF for A71. This means drivers interpreted what lies ahead of them from the state of road markings. Hence, transport engineers should work with materials of road markings in order to provide multi modal feedback to improve driver SA. Engineers can also focus on making these markings reflective when approaching an intersection and this could be very useful in night driving.

10.6.3 SER of different Road User Groups

A relationship is established between SER and SA in this thesis through CSF. It is these CSF which contribute towards forming the perceived image of the road to the driver. The car driver drives his/her vehicle forward on the basis of the information conveyed through CSF integrated in the road design and activating appropriate schemata. The CSF used in this thesis has been derived from the perspective of car drivers. Roads have all forms of road users in it; such as cyclists, motorcyclists, lorries and farm vehicles. Previous research has proved that each road user group (cars, motorcyclists and cyclists) interpret the same situation differently (Salmon et al., 2013, Walker et al., 2010). All road user groups have varied cognitive tasks (driving a car vs riding a bike) to perform and have different goals (Car drivers: Making sure what vehicles are in front and behind

them, what lies ahead, checking speed, etc. Cyclists: be very careful of parked cars on the road, look for alternative paths such as footpath and cycling lane) to achieve. This implies each road user group has a different perceived image of the road. Therefore, it will be wise to assume that each road user group would have different set of CSF from the other group. Hence, the next immediate strand of research is to explore if SER for one road user group is the same as for another. Are there any conflicting knowledge objects in amongst road users leading to potential accidents? This can be done to check to what extent the CSF used in this thesis is common across other road user groups and how are they integrated in the schema.

10.6.4 Extending SER

Since a relationship is associated between SER and SA in this thesis future research should now focus on making it cost and time effective. For instance a research question which has emerged from this thesis is whether it is possible to make an ordinary road self-explaining or to make it more user friendly. Chapter 4 has provided the SA requirements results which show that no one source of information can be over emphasized. Following a systems approach every source supplies SA to the driver. Therefore, the next stage of research should explore how an existing road can be improved by changing vehicle dynamics such as making more vulnerable road users more conspicuous, providing drivers with more information about what lies ahead of them through in-vehicle assistance systems and supplying them with more than one mode of feedback. Re-engineering of roads is expensive and time consuming hence, effectiveness of small interventions should be tested on driver behaviour and SA.

10.7 Closing Statement

Situation awareness is an established safety critical construct. User friendly roads can be designed by adopting a self-explaining approach to road safety. That is to say ensuring a road is homogeneous within the same categories, heterogeneous across different categories and predictable. The research findings show that distributed situation awareness is the best available model to design a user friendly road, or in other words, a SA rich system. The research shows empirically that an important relationship exists between SA and SER. Using the propositional network methodology, endemic features can be extracted and social network analysis used to sieve out the cognitive

salient features from the other unimportant road characteristics. The thesis ends with a road drivability tool. This is a potentially very powerful commodity which can predict the level of SA afforded by a particular system. These findings are very encouraging for transportation engineers. It is expected through publication of SER-SA research that the meaningfulness of the findings will be disseminated among other researchers.

The aim of this thesis has been to present a substantial original contribution to knowledge. This has been achieved in terms of how SA is viewed in road transportation; how the SER approach can be followed through with propositional networks, enabling a numerical value to be ascribed to SA features and possible measures to enhance it. The discovered relationship between SER and SA is a powerful one that can be advanced even further to the ultimate benefit of road safety, usability and performance.

References

AASTHO (1990). *A policy on geometric design of highways and streets*. Washington, D.C: American Association of State Highway and Transportation Officials (AASHTO).

Aarts, L.T. & Davidse, R.J**.** (2007). *Herkenbare vormgeving van wegen.* R-2006-18. SWOV, Leidschendam.

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, 37*(4), 445-456.

Ackerman, Robert K. "Army Intelligence Digitizes Situational Awareness." <u>Signal</u>. Armed Forces Communications and Electronics Association. 2005. Retrieved July 19, 2014 from High Beam Research: http://www.highbeam.com/doc/1P3-878237791.html

Adams, J. A. (2005). Human-Robot Interaction Design: Understanding User Needs and Requirements. *Human Factors and Ergonomics Society Annual Meeting Proceedings, 49*(3), 447-451.

Adams, J. A., Humphrey, C. M., Goodrich, M. A., Cooper, J. L., Morse, B. S., Engh, C., & Rasmussen, N. (2009). Cognitive task analysis for developing unmanned aerial vehicle wilderness search support. *Journal of cognitive engineering and decision making, 3*(1), 1-26.

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors, 37*(1), 85-104. doi: 10.1518/001872095779049462

Ajzen, I. (1991). The theory of planned behaviour. *Organizational behaviour and human decision processes, 50*(2), 179-211.

Alberts, D. S., Garstka, J. J., & Stein, F. P. (1999). *Network Centric Warfare: Developing and Leveraging Information Superiority*. Assistant Secretary of Defense (C3I/Command Control Research Program) Washington D.C.

Allen, T. M., Lumenfeld, H., & Alexander, G. J. (1971). *Driver information needs* (No. HS-011 999).

Anastasi, A. (1995). Psychology evolving: Linkages, hierarchies, and dimensions.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*: VH Winston & Sons.

Anderson, R. (1977). The notion of schemata and the educational enterprise *Schooling and the acquisition of knowledge B2 - Schooling and the acquisition of knowledge*. Hillsdale, N. J.: Erlbaum.

Andre, T. S., & Schopper, A. W. (1997). *Human Factors Engineering in System Design*: British Columbia Teacher.

Annett, J. (2002). A note on the validity and reliability of ergonomics methods. *Theoretical Issues in Ergonomics Science, 3*(2), 228-232. doi: 10.1080/14639220210124067

Annett, J. (2005). Hierarchical task analysis *Handbook of Human Factors and Ergonomics Methods B2 - Handbook of Human Factors and Ergonomics Methods* (pp. 33.31-33.37). London: Taylor & Francis.

Annett, J., & Duncan, K. D. (1967). Task analysis and training design. *Journal of Occupational Psychology*, 41, 211-221

Annett, J., & Kay, H. (2005). Knowledge of results and skilled performance'. *Ergonomics: Skill, displays, controls, and mental workload, 2*, 91.

Artman, H. (2000). Team situation assessment and information distribution. *Ergonomics, 43*(8), 1111-1128.

Artman, H., & Garbis, C. (1998). Team communication and coordination as distributed cognition. In *Proceedings of 9th Conference of Cognitive Ergonomics: Cognition and cooperation* (pp. 151-156).

Attitudes, iniciativa social de Audi.   Retrieved June 15, 2014, from http://www.tobii.com/fr/eye-tracking-research/france/bibliotheque/customer-cases/psychologie-cognitive/attitudes-iniciativa-social-de-audi/

Axelrod, R. (1973). Schema Theory: An Information Processing Model of Perception and Cognition. *The American Political Science Review, 67*(4), 1248-1266.

Baber, C. (2006). Cognitive aspects of tool use. *Applied ergonomics, 37*(1), 3-15.

Baber, C., & Mirza, M. G. (1998). Ergonomics and the evaluation of consumer products: surveys of evaluation practices. *Human Factors in Consumer Products*, 91-104.

Baber, C., & Stanton, N. A. (2002). Task analysis for error identification: theory, method and validation. *Theoretical Issues in Ergonomics Science, 3*(2), 212-227.

Bainbridge, L. (1999). Verbal reports as evidence of the process operator's knowledge. *International Journal of Human-Computer Studies, 51*(2), 213-238. doi: http://dx.doi.org/10.1006/ijhc.1979.0307

Banbury, & Tremblay. (2004). *A cognitive approach to situation awareness : theory and application / edited by Simon Banbury and Sebastien Tremblay*: Ashgate, 2004.

Banks, A. P., & Millward, L. J. (2009). Distributed Mental Models: Mental Models in Distributed Cognitive Systems. *Journal of Mind and Behaviour, 30*(4), 249-266.

Banks, V. A., Stanton, N. A., & Harvey, C. (2014). What the drivers do and do not tell you: using verbal protocol analysis to investigate driver behaviour in emergency situations. *Ergonomics, 57*(3), 332-342. doi: 10.1080/00140139.2014.884245

Bartlett, F. C., & Kintsch, W. (1995). *Remembering: A Study in Experimental and Social Psychology.* Cambridge: Cambridge University Press, 1995.

Baumann, M., & Krems, J. F. (2007). Situation awareness and driving: A cognitive model *Modelling driver behaviour in automotive environments* (pp. 253-265): Springer.

Bedny, G., & Meister, D. (1999). Theory of activity and situation awareness. *International Journal of cognitive ergonomics*, *3*(1), 63-72.

Behavioural Change Theories. (2014). from http://www.roadsafetyevaluation.com/evaluationtopics/interventionplanning/behavioural-change-theories.html

Bekiaris, E., Amditis, A., & Panou, M. (2003). DRIVABILITY: a new concept for modelling driving performance. *Cognition, Technology & Work, 5*(2), 152-161. doi: 10.1007/s10111-003-0119-x

Bekiaris, E., & Gaitanidou, E. (2011). Towards forgiving and self-explanatory roads *Infrastructure and Safety in a Collaborative World* (pp. 15-22): Springer Berlin Heidelberg.

Bell, H. H., & Lyon, D. R. (2000). Using observer ratings to assess situation awareness. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement.* (pp. 129-146). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88*(4), 354-364. doi: 10.1037/0033-295x.88.4.354

Berndt, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81*, 211-241. doi: 10.1016/0001-6918(92)90019-a

Billings, C. E. (1995). Situation awareness measurement and analysis: A commentary *Proceedings of the International Conference on Experimental Analysis and Measurement of Situation Awareness*. FL: Embry-Riddle Aeronautical University Press.

Bolia, R. S., Vidulich, M. A., Nelson, W. T., & Cook, M. J. (2006). A history lesson on the use of technology to support military decision-making and command & control. *Decision making in complex systems* (pp. 191-200). Aldershot, UK: Ashgate.

Box, E. (2011). Mortality Statistics and road Traffic Accidents in U.K: RAC Foundation.

Brannick, M. T., Salas, E., & Prince, C. W. (Eds.). (1997). *Team performance assessment and measurement: Theory, methods, and applications*. Psychology Press

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta psychologica, 81*(3), 211-241.

Brewer, M. B. (2000). Research design and issues of validity. *Handbook of research methods in social and personality psychology*, 3-16.

Brewer, W. F., & Nakamura, G. V. (1984). The Nature and Functions of Schemas. Technical Report No. 325.

Brian, H., Carmit, K. M., & Paul, N. G. (2007). Distributed cognition in the heart room: How situation awareness arises from coordinated communications during cardiac surgery. *Journal of Biomedical Informatics, 40*, 539-551. doi: 10.1016/j.jbi.2007.02.001

Brumec, U., Herman, E., Hrabar, N. V., & Polič, M. (2010). *Road Environment as a Factor of Road Safety: Do we understand the concept?''*. Paper presented at the Road Accident Prevention 2010, Novi Sad, Serbia. Retrieved from http://www.ncpiarc.si/dokumenti/gradivo_clanov/Brumec_clanek1_Road%20environment.pdf on June 29, 2013

Brunswik, E. (1952). The conceptual framework of psychology. (Int. Encycl. unified Sci., v. 1, no. 10.).

Calvert, M., & Walker, G. (2013). Modelling Driver Behaviour at Roadworks. Draft: Transport Scotland.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan, Jr. (Ed.), *Individual and group decision making: Current issues.* (pp. 221-246). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Carol, L. (1992). Desperately seeking SA. *TAC attack, 32*(TAC SP 127-1), 5-6.

Carsten, O., Kircher, K., & Jamson, S. (2013). Vehicle-based studies of driving in the real world: The hard truth? *Accident Analysis & Prevention*, *58*, 162-174.

Charles, C. M. (1998). *Introduction to educational research*: ERIC.

Charlton, S. G., Mackie, H. W., Baas, P. H., Hay, K., Menezes, M., & Dixon, C. (2010). Using endemic road features to create self-explaining roads and reduce vehicle speeds. *Accident Analysis waand Prevention, 42*(6), 1989-1998. doi: 10.1016/j.aap.2010.06.006

Chase, V. M. (1998). Sources of Power: How People Make Decisions, G Klein. *Nature-London-*(6673), 242.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive psychology, 4*(1), 55-81.

Chimir, I. A., Abu-Dawwas, W. A., & Horney, M. A. (2005). Neisser's Cycle of Perception: Formal Representation and Practical Implementation. *Journal of Computer Science, 1*, 106.

Cocu, X., Casse, C., Volckaert, A., & Van Geem, C. (2011). Self-explaining road treatments:Report from expert workshop *Speed Adaption Control by Self-Explaining Roads* (Vol. 3, pp. 28): Belgian Road Research Centre.

Cohen, J. (2013). *Statistical power analysis for the behavioural sciences*: Routledge Academic.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review, 82*(6), 407.

Collins, A. M., & Quillian, M. R. (1972). Experiments on semantic memory and language comprehension.

Coombs, M. (2010). Conciliatory Modelling of Web Sources Using Automated Content Analysis: A Vignette on Drilling for Shale Gas in Upstate New York. Retrieved June 20th, 2014, from https://www.Leximancer™ .com/lmedia/Facilitating_Open_Government.pdf

ETSC (1995) Reducing traffic injuries resulting from excess and inappropriate speed. Brussels: European Transport Safety Council. Retrieved from http://archive.etsc.eu/documents/Reducing%20traffic%20injuries%20from%20excess%20and%20inappropriate%20speed.pdf on January 10, 2014.

Coyne, P. (2007). Roadcraft: the essential police driver's handbook.

Cretchley, J., Rooney, D., & Gallois, C. (2010). Mapping a 40-year history with Leximancer™: Themes and concepts in the Journal of Cross-Cultural Psychology. *Journal of Cross-Cultural Psychology, 41*(3), 318-328.

Cronbach, L. J. (1954). Report on a psychometric mission to Clinicia. *Psychometrika, 19*(4), 263-270.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, *52*(4), 281.

CROW (2004). Richtlijn essentiële herkenbaarheidkenmerken van weginfrastructuur: wegwijzer voor implementatie. [Directive essential recognizability characteristics for infrastructure: manual for implementation] (In Dutch). Publicatie 203. CROW, Ede.

De Waard, D. (1996). *The measurement of drivers' mental workload*. Groningen University, Traffic Research Centre.

Dekker, S., Cook, R., & Sarter, N. (2010). *Behind human error*. Farnham: Ashgate.

Dennehy, K. (1997). Cranfield situation awareness scale.

Department of Transport, V. (2000). UPDATE Virginia Keeps America Moving. *Variance in Speed, Not Speed Limit Poses Greatest Risk on Interstate Roads*(4). Retrieved from http://www.virginiadot.org/projects/resources/newsl-I81-winter00%20.pdf on April 20, 2014.

Diaper, D. (1989). *Knowledge elicitation: principles, techniques and applications / by Dan Diaper*: Ellis Horwood, 1989.

Diaper, D. (2001). Task analysis for knowledge descriptions (TAKD): a requiem for a method. *Behaviour & Information Technology, 20*(3), 199-212.

Diaper, D., & Stanton, N. A. (2004). *The handbook of task analysis for human-computer interaction*. Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., et al (2006). The 100-car naturalistic driving study, phase ii – Results of the 100-car field experiment DOT HS 810 593. Washington, DC: National Highway Traffic Safety Administration

Doerner, D. (1980a). On the Difficulties People Have in Dealing with Complexity. *Simulation & Gaming, 11*(1), 87-106. doi: 10.1177/104687818001100108

Doerner, D. (1980b). On the Difficulties People have in Dealing with Complexity. *Simulation & Gaming, 11*(1), 87.

Durso, F. T. (1999). *Handbook of applied cognition*. New York, NY US: John Wiley & Sons Ltd.

Eason, K. (1988). *Information Technology And Organisational Change*: Taylor & Francis.

Edgar, G. K., & Edgar, H. E. (2007). Using signal detection theory to measure situation awareness: the technique, the tool (QUASA), the test, the way forward *Decision making in complex environments* (pp. 373-385). Aldershot, UK: Ashgate.

Edquist, J., Rudin-Brown, C. M., & Lenne, M. (2009). Road Design Factors and their Interactions with Speed and Speed Limits (pp. 30). Victoria, Australia: Monash University Accident Research Centre.

Eenink, R., Barnard, Y., Baumann, M., Augros, X., & Utesch, F. (2014). *UDRIVE: the European naturalistic driving study*. Paper presented at the Transport Research Arena, Paris. Retrieved from http://www.traconference.eu/papers/pdfs/TRA2014_Fpaper_19523.pdf on June, 6, 2014.

Elliott, M. A., McColl, V., Kennedy, J. V., & Department for, T. (2003). *Road design measures to reduce drivers' speed via 'psychological' processes*.

Endsley, M. R., & Garland, D. J. (Eds.). (2000). *Situation awareness analysis and measurement*. CRC Press.

Endsley, M. R. (1993). A Survey of Situation Awareness Requirements in Air-to-Air Combat Fighters. *International Journal of Aviation Psychology, 3*(2), 157.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems (Vol. 37, pp. 32-64).

Endsley, M. R. (1999). Situation Awareness in Aviation Systems .*Handbook of Aviation Human Factors*. 257-276.

Endsley, M. R. (2000). Theoretical underpinning of situation awareness: a critical review. *Situation awareness analyses and measurement* (pp. 220-245). Mahawh, NL: Lawrence Erlbaum Publishers.

Endsley, M. R., Bolte, B., & Jones, D. G. (2003*). Designing for situation awareness: An approach to user centred design.* CRC Press.

Endsley, M. R., & Garland, D. J. (2000, July). Pilot situation awareness training in general aviation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, No. 11, pp. 357-360). SAGE Publications.

Endsley, M. R., & Garland, D. J. (2000). *Situation awareness analysis and measurement*: CRC Press.

Endsley, M. R., & Jones, W. (2013). Situation awareness. *The Oxford Handbook of Cognitive Engineering*, 88-108.

Endsley, M., & Jones, W. M. (1997). *Situation Awareness Information Dominance & Information Warfare*. Logicon technical Services Inc. Dayton OH.

Endsley, M. R., & Jones, W. M. (2001). A model of inter- and intra-team situation awareness: Implications for design, training and measurement *New trends in*

cooperative activities: Understanding system dynamics in complex environments B2. Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R., Robertson, M. M., & Porter, K. (1999). *Evaluation of A Program For Training Skills Related To Team Situation Awareness In Aviation Maintenance*.

Endsley, M. R., & Rodgers, M. D. (1997). *Distribution of attention, situation awareness, and workload in a passive air traffic control task [microform]: implications for operational errors and automation*. Washington, D.C.: U.S. Dept. of Transportation, Federal Aviation Administration, Office of Aviation Medicine ; Springfield, Va. : Available to the public through the National Technical Information Service, [1997].

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). *A comparative analysis of SAGAT and SART for evaluations of situation awareness.* Paper presented at the Proceedings of the human factors and ergonomics society annual meeting.

Environmental Systems Research Institute (ESRI) (2008). Public Safety and Homeland Security: Situation Awareness. Retrieved from http://www.esri.com/library/whitepapers/pdfs/situational-awareness.pdf. On March 20, 2012.

European Transport Safety Council (ETSC) (1995). Reducing Traffic Injuries Resulting from Inappropriate and Excessive Speed. Brussels, Belgium.40.

Evans, S. H., & Arnoult, M. D. (1967). Schematic concept formation: Demonstration in a free sorting task (Vol. 9).

Evans, S. H., & Edmonds, E. M. (1966). Schema Discrimination as a Function of Training (Vol. 5).

Flach, J. M. (1995). Situation awareness: proceed with caution (Vol. 37, pp. 149-157).

Forrester, J. W. (1961). Industrial Dynamics *Industrial Dynamics B2 - Industrial Dynamics*. Cambridge, MA: Massachusetts Institute of Technology.

Forrester, J. W. (1975). Collected Papers of Jay W. Forrester Cambridge, MA: Wright Allen Press.

Forrester, J. W. (1995). The beginning of system dynamics. *McKinsey Quarterly*, 4-17.

Fracker, M. L. (1991). Measures of situation awareness: Review and future directions: DTIC Document.

French, D. J., West, R. J., Elander, J., & Wilding, J. M. (1993). Decision-making style, driving style, and self-reported involvement in road traffic accidents. *Ergonomics, 36*(6), 627-644.

Garber, N. J., & Gadiraju, R. (1989). Factors Affecting Speed Variance and its Influence on Accidents. *Transportation Research Record*(1213), p. 64-71.

Gibson, J. J. (1977). The theory of affordances. *Hilldale, USA*.

Gibson, J. J. (2013). *The ecological approach to visual perception*: Psychology Press.

Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory, 6*(3), 225-255.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The qualitative report, 8*(4), 597-607.

Goldenbeld, C., & Van Schagen, I. (2007). The credibility of speed limits on 80km/h rural roads: The effects of road and personality characteristics. *Accident Analysis and Prevention*(6), 1121. doi: 10.1016/j.aap.2007.02.012

Goldsmith, M. (2002). Try Feedforward Instead of Feedback. Retrieved from
http://www.marshallgoldsmithlibrary.com/cim/articles_display.php?aid=110 on
January 15, 2014

Golightly, D., Wilson, J. R., Lowe, E., & Sharples, S. (2010). The role of situation
awareness for understanding signalling and control in rail operations.
Theoretical Issues in Ergonomics Science, 11(1/2), 84-98. doi:
10.1080/14639220903009961

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation
awareness in decentralized command and control environments. *Ergonomics,
49*, 1312.

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation
awareness in decentralized command and control environments. *Ergonomics,
49*(12-13), 1312-1325. doi: 10.1080/00140130600612788

Green, M. (2009). Self-explaining roads. *Nordic Road and Transport Research* (3).

Gregory, R. L. (1970). The intelligent eye. London: Weidenfeld & Nicolson

Groner, N. (2011). Can the Cognitive Engineering Approach Prevent "Normal
Accidents"? How Design Might Improve Societal Resiliency to Critical
Incidents. *Journal of Critical Incident Analysis, 1*(2), 9.

Groner, N. E. (2009). Situation Awareness Requirement Analysis for Emergency
Management Planning. Retrieved Working Paper [09-02], from
http://christianregenhardcenter.org/pdfs/Situation_Awareness_Requirements.pdf
on February 10, 2014

Groth-Marnat, G. (2009). *Handbook of psychological assessment*: John Wiley & Sons.
4th Edition

Gugerty, L. (1998). Evidence from a partial report task for forgetting in dynamic spatial memory. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 40*(3), 498-508.

Gugerty, L. (2011). Situation awareness in driving. *Handbook for driving simulation in engineering, medicine and psychology*.

Gugerty, L. J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied, 3*(1), 42-66. doi: 10.1037/1076-898x.3.1.42

Haines, R. F., & Courtney, F. (1992). *Night Flying*. New York: USA: McGraw-Hill Education.

Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Halverson, C. A. (1995). *Inside the cognitive workplace: new technology and air traffic control* (Doctoral dissertation). University of California. San Diego.

Hamilton, I. W., Lowe, E., & Hill, C. (2007). Early route drivability assessment in support of railway investment. *People and Rail Systems: Human Factors at the Heart of the Railway*, 185.

Hamilton, K., & Kennedy, J. (2005). Rural Road Safety: A literature Review. In S. Research (Ed.).

Hamilton, W. (1997). *Situation Awareness* Metrics Program. *SAE Technical Paper*. doi: 10.4271/871767

Hanson, M. (2003). *Contemporary ergonomics 2001*: CRC Press.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology, 52*, 139-183.

Hartel, C., Smith, K. A., & Prince, C. (1991). Defining aircrew coordination: Searching mishaps for meaning *Sixth International Symposium of Aviation Psychology B2 - Sixth International Symposium of Aviation Psychology*. Columbus, OH.

Hastie, R., & Park, B. (1986). The Relationship Between Memory and Judgment Depends on Whether the Judgment Task is Memory-Based or On-Line. *Psychological Review, 93*(3), 258-268.

Head, H. (1920). *Studies in Neurology* (Vol. 1). London: Oxford University Press.

Held, C., Knauff, M., & Vosgerau, G. (2006). *Mental models and the mind: Current developments in cognitive psychology, neuroscience, and philosophy of mind*. Amsterdam Netherlands: Elsevier.

Hendy, K. (1995). *Situation Awareness and Workload: Birds of a Feather?* : DTIC Document.

Herrstedt, L. (2006). *Self-explaining and Forgiving Roads–Speed management in rural areas.* Paper presented at the Australian Road Research Board (ARRB) Conference.

Hewett, D. G., Watson, B. M., Gallois, C., Ward, M., & Leggett, B. A. (2009). Communication in medical records intergroup language and patient care. *Journal of Language and Social Psychology, 28*(2), 119-138.

Hickman, J.S., Hanowski, R.J., 2010. Evaluating the safety benefits of a low cost driving behavior management system in commercial vehicle operations. Report No.FMCSA-RRR-10-033. Federal Motor Carrier Safety Administration, U. S. Department of Transportation, Washington, DC.

Hodges, N. J., Hayes, S. J., Breslin, G., & Williams, M. A. (2005). An evaluation of the minimal constraining information during observation for movement reproduction. *Acta Psychologica, 119*, 264-282. doi: 10.1016/j.actpsy.2005.02.002

Hole, G. (2007). *The psychology of driving*. Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*: CRC Press.

Horswill, M. S., & Coster, M. E. (2002). The effect of vehicle characteristics on drivers' risk-taking behaviour. *Ergonomics, 45*(2), 85-104.

Horswill, M. S., & McKenna, F. P. (1999). The development, validation, and application of a video-based technique for measuring an everyday risk-taking behaviour: Drivers' speed choice. *Journal of Applied Psychology, 84*(6), 977-985. doi: 10.1037/0021-9010.84.6.977

Humphrey, C. M. (2009). *Information abstraction visualization for human-robot interaction.* Vanderbilt University.

Humphrey, C. M., & Adams, J. A. (2009). Robotic tasks for chemical, biological, radiological, nuclear and explosive incident response. *Advanced Robotics, 23*(9), 1217-1232.

Humphrey, C. M., & Adams, J. A. (2010). General visualization abstraction algorithm for directable interfaces: Component performance and learning effects. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 40*(6), 1156-1167.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA US: The MIT Press.

Hutchins, E. (2000). Distributed Cognition. International Encyclopaedia of the Social and Behavioural Sciences (IESBS).

Hutchins, E., & Klausen, T. (1996). Distributed Cognition in an Airline Cockpit. *Cognition and communication at work*, 15-34.

Isaac, A. R. (1997). Situational awareness in air traffic control human cognition and advanced technology *Engineering Psychology and Cognitive Ergonomics B2 - Engineering Psychology and Cognitive Ergonomics*. Aldershot, UK: Ashgate.

James, L. (1997). Principles of Driving Psychology.   Retrieved March 10, 2012, from http://drdriving.org/articles/principles.htm

Jenkins, D. P., Stanton, N. A., Salmon, P. M., Walker, G. H., & Young, M. S. (2008). Using cognitive work analysis to explore activity allocation within military domains. *Ergonomics, 51*(6), 798-815. doi: 10.1080/00140130801915246

John, D. S. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43*, 301-335. doi: 10.1016/0749-5978(89)90041-1

Johnson, C. W. (2006). What are emergent properties and how do they affect the engineering of complex systems? *Reliability Engineering & System Safety, 91*(12), 1475-1481.

Johnson-Laird, P. N. Mental models: Towards a cognitive science of language, inference, and consciousness / P.N. Johnson-Laird: Cambridge, EUA : Harvard University 1983.

Jon, G. H., & Andrés, S. (2008). A conceptual model for the analysis of mishaps in human-operated safety-critical systems. *Safety Science, 46*, 22-37. doi: 10.1016/j.ssci.2006.11.001

Jordan, P. W., Thomas, B., McClelland, I. L., & Weerdmeester, B. (1996). *Usability evaluation in industry*: CRC Press.

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science, 5*(2), 113-153. doi: 10.1080/1463922021000054335

Kanis, H. (2000). Questioning validity in the area of ergonomics/human factors. *Ergonomics, 43*(12), 1947-1965.

Kaptein, N. (1998). *Effects of cognitive road classification on driving behaviour: a driving simulator study.* University College London.

Kaptein, N., Janssen, W., & Claessens, M. (2002). A study of subjective road categorization and driving behaviour. *Human Factors for Highway Engineers*.

Karwowski, W. (2001). *International Encyclopaedia of Ergonomics and Human Factors*: Taylor & Francis.

Kirk, J., & Miller, M. L. (1988). Reliability and Validity in Qualitative Research. *International Journal of Qualitative Studies in Education, 1*(1).

Klauer, S. G., Perez, M., & McClafferty, J. (2011). Naturalistic driving studies and data coding and analysis techniques. *Handbook of Traffic Psychology*, 73-85.

Kumar, J., Yammiyavar, P., & Nielsen, J. (2007). Mind Tape technique—A usability evaluation method for tracing cognitive processes in cross cultural settings. *eMinds*, *1*(3), 83-99.

Lerner, N. D. (2003). *Additional investigations on driver information overload* (No. 488). Transportation Research Board.

Lissitz, R. W. (2009). *The concept of validity: Revisions, new directions, and applications*: IAP.

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30*(5), 583-598.

Macadam, C. (2003). Understanding and modelling the human driver. *Vehicle System Dynamics, 40*(1-3), 101-134.

Macharis, C., Verbeke, A., Brucker, K., Gelova, E., Weinberger, J & Vasek, J. (2005). Implementation scenarios and further research priorities regarding forgiving and self-explaining roads. Deliverable 5.3. In-Safety Project.

Mackie, H. W., Charlton, S. G., Baas, P. H., & Villasenor, P. C. (2013). Road user behaviour changes following a self-explaining roads intervention. *Accident Analysis & Prevention, 50*, 742-750. doi: 10.1016/j.aap.2012.06.026

Malaterre, G. (1990). Error analysis and in-depth accident studies. *Ergonomics, 33*(10-11), 1403-1421. doi: 10.1080/00140139008925341

Mandler, J. M. (1984). *Stories, Scripts, and Scenes: Aspects of Schema Theory*.

Mares, P., & Coyne, P. (2013). *Roadcraft: The Police Drivers Handbook* (New 2013 ed.). London: Police Foundation.

Marshall, S. P. (1995). *Schemas in Problem Solving / Sandra P. Marshall. [eBook]*: Cambridge : Cambridge University Press, 1995.

Martens, H., Compte, S., & Kaptein, N. A. (1997). *The Effects of Road Design on Speed Behaviour: A Literature Review*: TNO Human Factors Research Institute.

Martens, M. H. (2007). *The Failure to Act Upon Important Information: where Do Things Go Wrong? (*Doctoral Dissertation*).* Vrije Universiteit. Amsterdam

Matena, S., Weber, R., Louwerse, R., Drolenga, H., Vaneerdewegh, P., Pokorny, P., .Hollo, P. (2006). Road Categorisation and Design of Self- Explaining Roads *Road Infrastructure Safety Protection – Core-Research and Development for Road Safety in Europe; Increasing safety and reliability of secondary roads for a sustainable Surface Transport (RIPCORD - ISEREST)* (pp. 132). Belgium.

Matthews, M. D., Strater, L. D., & Endsley, M. R. (2004). Situation Awareness Requirements for Infantry Platoon Leaders. *Military Psychology, 16*(3), 149-161.

Mazet, C., Dubois, D., & Fleury, D. (1987). Catégorisation et interprétation de scènes visuelles: le cas de l'environnement urbain et routier. *Psychologie française, 32*(1-2), 85-95.

McGuinness, B., & Foy, L. (2000). *A Subjective Measure of SA: The Crew Awareness Rating Scale (CARS).*

McKnight, A. J., Adams, B. B., & Human Resources Research Organization, A. V. A. (1970). *Driver Education Task Analysis. Volume II: Task Analysis Methods. Final Report.*

Meadows, D. L. (1974). *Dynamics of growth in a finite world*: Wright-Allen Press.

Meister, D. (1985). *Behavioural analysis and measurement methods.* Wiley, 1985.

Michon, J. A. (1993). *Generic Intelligent Driver Support*: Taylor & Francis.

Minsky, M. L., & Papert, S. A. (1987). *Perceptrons - Expanded Edition: An Introduction to Computational Geometry*. Boston, MA: MIT press.

Moray, N. (2004). *Ou' sont les neiges d' antan?* Paper presented at the Human Performance, Situation Awareness and Automation: Current Research and Trends: HPSAA II, Daytona Beach, Florida.

Mosier, C. I. (1947). A critical examination of the concepts of face validity. *Educational and Psychological Measurement*.

Munduteguy, C., & Ragot-Court, I. (2011). A Contribution to Situation Awareness Analysis: Understanding How Mismatched Expectations Affect Road Safety. *Human Factors, 53*(6), 687-702.

Mundutéguy, C. (2011). A Contribution to Situation Awareness Analysis Understanding How Mismatched Expectations Affect Road Safety. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(6), 687-702.

Myers, A. M., Trang, A., & Crizzle, A. M. (2011). Naturalistic study of winter driving practices by older men and women: examination of weather, road conditions, trip purposes and comfort. *Canadian Journal on Aging, 30*(4), 577-589.

National Highway Traffic Safety Administration (NHTSA) (1991). Effect of Car size on fatality and Injury Risk. Washington D.C.

Karwoski, W. (2001). Core Competencies in Ergonomics in *International Encyclopaedia of Ergonomics and Human Factors* London: Taylor & Francis.

Neisser, U. Cognition and reality: principles and implications of cognitive psychology. San Francisco, EUA: W.H. Freeman and Company 1976.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton Century Crofts.

Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement, 22*(4), 287-293.

Nielsen, J. (1992). The Usability Engineering Life-cycle. *Computer, 25*(3), 12-22.

Niu, L., Lu, J., & Zhang, G. (2009). *Cognition-driven decision support for business intelligence : models, techniques, systems and applications.* Berlin ; Heidelberg : Springer, c2009.

Norman.D. (2002). *The design of everyday things / by Donald A. Norman*: Basic Books, 2002.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88*(1), 1-15. doi: 10.1037/0033-295x.88.1.1

Norman, D. A. (1986). Cognitive engineering. *User Centred System Design*, 31-61.

Norman, D. A. (1999). Affordance, conventions, and design. *interactions, 6*(3), 38-43. doi: 10.1145/301153.301168

Oden, G. C. (1987). Concept, Knowledge and Thought. *Annual Review of Psychology, 38*(1), 203.

Olarte, O. (2011, April, 2011). Human Error accounts for 90% of Road Accidents. *AlertDriving Magazine*.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. L. (2009). Driver distraction in commercial vehicle operations.

Ottino, J. M. (2003). Complex systems. *AIChE Journal, 49*(2), 292-299. doi: 10.1002/aic.690490202

Pass, F., Renkl, A., & Sweller, J. (2004). Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science, 32*(1-2), 1-8.

Patricia, W. C., Keith, J. H., Richard, E. N., & Lindsay, M. O. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18*, 293-328. doi: 10.1016/0010-0285(86)90002-2

Patrick, J., Gregov, A., & Halliday, P. (2000). Analysing and training task analysis. *Instructional science, 28*(1), 51-79.

Patrick, J., & Morgan, P. L. (2010). Approaches to understanding, analysing and developing situation awareness. *Theoretical Issues in Ergonomics Science, 11*(1/2), 41-57. doi: 10.1080/14639220903009946

Paul, P., & Mukhopadhyay, K. (2004). Verbal Protocol Analysis *Encyclopaedia of Social Science Research Methods*: Sage Publications.

Perla, P. P., Nogfi, A., Loughran, J., & Stahl, M. (2000). Gaming and Shared Situation Awareness *Gaming and Shared Situation Awareness B2 - Gaming and Shared Situation Awareness*. Alexandria, VA: Centre for Naval Analyses.

Piaget, J. (1929). *The Child's Conception of the World*. London: Routledge and Kegan Paul, Ltd.

Pisarski, A. E. (1986). DEEP-SIX 55. *Reason, 17*(6), p. 32-35.

Plant, K. L., & Stanton, N. A. (2012). Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model. *Safety Science*. doi: 10.1016/j.ssci.2011.09.005

Posner, M. I., & Keele, S. W. (1968). On the Genesis of Abstract Ideas. *Journal of Experimental Psychology, 77*(31), 353-363. doi: 10.1037/h0025953

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83*(21), 304-308. doi: 10.1037/h0028558

Pritchett, A. R., Hansman, R. J., & Johnson, E. N. (1995). Use of testable responses for performance-based measurement of situation awareness. *Experimental Analysis and Measurement of Situation Awareness* (pp. 75-81). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Proctor, R. W., & Van Zandt, T. (2008). *Human Factors in Simple and Complex Systems, Second Edition*: Taylor & Francis.

Quillan, M. R. (1966). Semantic memory: DTIC Document.

Rafferty, L. A., Stanton, N. A., & Walker, G. H. (2013). Great expectations: A thematic analysis of situation awareness in fratricide. *Safety Science*, 63. doi: 10.1016/j.ssci.2012.03.020

Regan, M., Williamson, A., Grzebieta, R., Charltonb, J., Lenneb, M., Watsonc, B., Anderson, R. (2013). *The Australian 400-car Naturalistic Driving Study: Innovation in road safety research and policy.* Paper presented at the Australasian Road Safety Research Policing Education Conference, 2013, Brisbane, Queensland, Australia.

Regan, M., Williamson, A., Grzebieta, R., & Tao, L. (2012). *Naturalistic driving studies: literature review and planning for the Australian naturalistic driving study.* Paper presented at the Australasian College of Road Safety Conference 2012, Sydney, New South Wales, Australia.

Rogers, Y. (1992). Ghosts in the network: distributed troubleshooting in a shared working environment. In Proceedings of the 1992 ACM conference on Computer-supported cooperative work (pp. 346-355). ACM.

Rogers, Y. (1993). Coordinating Computer-Mediated Work. *Computer Supported Cooperative Work (CSCW)*, *1*(4), 295-315.

Rogers, Y. (1997). A brief introduction to distributed cognition. *Retrieved July*, *24*, 2012.

Rooney, D., Paulsen, N., Callan, V. J., Brabant, M., Gallois, C., & Jones, E. (2010). A new role for place identity in managing organizational change. *Management Communication Quarterly, 24*(1), 44-73.

Rouse, W. B., & Morris, N. M. (1985). On Looking into the Blackbox: Prospects and Limits in the Search for Mental Models (DTIC#AD-A159080) Atlanta, GA: Centre for Man-Machine Systems Research, Georgia Institute of Technology.

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design, and conduct effective tests*: John Wiley & Sons.

Rumelhart, D. E. (1984). Schemata and the cognitive system. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition, Vol 1.* (pp. 161-188). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Rumelhart, D. E., & Ortoney, A. (1977). The Representation of Knowledge in Memory *Schooling and the Acquisition of Knowledge B2 - Schooling and the Acquisition of Knowledge*. Hillsdale NJ: Erlbaum.

Rust, J., & Golombok, S. (1999). *Modern psychometrics: The science of psychological assessment*: Psychology Press.

Sabey, B. E., & Taylor, H. (1980). The known risks we run: the highway (pp. 43-70). Springer US.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors, 37*(1), 123-136. doi: 10.1518/001872095779049525

Salas, E., & Cannon-Bowers, J. A. (1997). Methods, Tools, and Strategies for Team Training. In Quinones, M. A., & Ehrenstein, A. E. (1997). *Training for a rapidly changing workplace: Applications of psychological research*. American Psychological Association.

Salmon, P.M., Young, K. L., Lenne, M. G., & Williamson, A. (Submitted). An on-road multi-method framework for examining driving errors and their causes. *Safety Science*.

Salmon, P. M. (2008). *Distributed situation awareness: Advances in theory, measurement and application to team work.* (Doctoral Dissertation), Brunel University, Middlesex

Salmon, P. M., Lenne, M. G., Walker, G. H., & Filtness, A. (2013). *Awesome foursome? the compatibility of driver, cyclist, motorcyclist, and pedestrian situation awareness at intersections.* Paper presented at the Proceedings of the 10th international conference on Engineering Psychology and Cognitive Ergonomics: applications and services - Volume Part II, Las Vegas, NV.

Salmon, P. M., Lenne, 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.

Salmon, P., Jenkins, D., Stanton, N., & Walker, G. (2010). Hierarchical task analysis vs. cognitive work analysis: comparison of theory, methodology and contribution to system design. *Theoretical Issues in Ergonomics Science*, *11*(6), 504-531.

Salmon, P. M., Read, G. J., Lenné, M. G., & Stanton, N. A. (2012). Through a different lens: applying contemporary systems analysis methods in road safety. In *Australasian Road Safety Research Policing Education Conference, 2012, Wellington, New Zealand*.

Salmon, P. M., Read, G. J., Stanton, N. A., & Lenné, M. G. (2013). The crash at Kerang: Investigating systemic and psychological factors leading to unintentional non-compliance at rail level crossings. *Accident Analysis & Prevention*, *50*, 1278-1288.

Salmon, P. M., Lenne, M. G., Walker, G. H., Stanton, N. A., & Filtness, A. (2014). Exploring schema-driven differences in situation awareness between road users: an on-road study of driver, cyclist and motorcyclist situation awareness. *Ergonomics, 57*(2), 191-209. doi: 10.1080/00140139.2013.867077

Salmon, P. M., & Stanton, N. A. (2013). Situation awareness and safety: Contribution or confusion? Situation awareness and safety editorial. *Safety Science, 56*, 1-5.

Salmon, P. M., Stanton, N. A., Gibbon, A., Jenkins, D., & Walker, G. H. (2010). *Human Factors Methods and Sports Science: A Practical Guide*. Florida, U.S.A: Taylor & Francis.

Salmon, P. M., Stanton, N. A., Jenkins, D. P., Walker, G. H., Young, M. S., & Aujla, A. (2007). What really is going on? Review, critique and extension of situation awareness theory *Engineering Psychology and Cognitive Ergonomics* (pp. 407-416): Springer.

Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science,9*, 297.

Salmon, P. M., Stanton, N. A., Walker, G. H., & Jenkins, D. (2009). *Distributed situation awareness: theory, measurement and application to teamwork* Ashgate.

Salmon, P. M., Stanton, N. A., & Young, K. L. (2011). Situation awareness on the road: review, theoretical and methodological issues, and future directions. *Theoretical Issues in Ergonomics Science*, 1-21. doi: 10.1080/1463922x.2010.539289

Salmon, Paul M., Lenné, Michael G., Walker, Guy H., & Filtness, Ashleigh J. (2013) Complex cognitive interactions in a badly designed world: Investigating the underlying causes of collisions between distinct road users. In *Australasian Road Safety Research, Policing and Education Conference 2013*, 28-30 August 2013, Brisbane Convention and Exhibition Centre, Brisbane, QLD.

Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied ergonomics*, *37*(2), 225-238.

Salvendy, G. (2012). *Handbook of human factors and ergonomics*: John Wiley & Sons.

Sanders, M. S., & McCormick, E. J. (1987). *Human factors in engineering and design* . McGraw-Hill Book Company.

Sandom, C. W. (2000). *Situation Awareness and Interactive System Safety Analysis.* (Doctoral Dissertation), Brunel University, Middlesex.

Sarter, L. D., Endsley, M.R., Pleban, R.J & Matthews,M.D. (2001). Measures of Platoon Leader Situation Awareness in Virtual Decision-Making Exercises (pp. 25). Alexandria, USA: US Army research institute for the behavioural and social sciences.

Sarter, N. B., & Woods, D. (1991). Situation Awareness: A Critical But ill-Defined Phenomenon. *International Journal of Aviation Psychology, 1*(1), 45.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37*(1), 5-19. doi: 10.1518/001872095779049516

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225-260.

Schuck, J. R., & Leahy, W. R. (1966). A comparison of verbal and non-verbal reports of fragmenting visual images. *Perception & Psychophysics*, *1*(3), 191-192.

Selcon, S., Taylor, R., & Koritsas, E. (1991). *Workload or situational awareness? TLX vs. SART for aerospace systems design evaluation.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Selcon, S. J., & Taylor, R. M. (1990). *Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design.* Paper presented at the {Situational Awareness in Aerospace Operations (AGARD-CP-478)}.

Senge, P. M., & Suzuki, J. (1994). *The fifth discipline: The art and practice of the learning organization* (p. 14). New York: Currency Doubleday.

Shadbolt, N. R., & Burton, M. (1995). Knowledge elicitation: a systemic approach (pp. 406-440).

Shiffrin, R. M., & Schneider, W. (1984). Automatic and controlled processing revisited.

Shu, Y., & Furuta, K. (2005). An inference method of team situation awareness based on mutual awareness. *Cognition, Technology & Work, 7*(4), 272.

Smith, A. E. (2003). *Automatic extraction of semantic networks from text using Leximancer™* . Paper presented at the Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology: Demonstrations-Volume 4.

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Hum. Factors, 37*, 137.

Sorensen, L. J., & Stanton, N. A. (2011). Is SA shared or distributed in team work? An exploratory study in an intelligence analysis task. *International Journal of Industrial Ergonomics*, *41*(6), 677-687.

Sorensen, L. J., Stanton, N. A., & Banks, A. P. (2011). Back to SA school: contrasting three approaches to situation awareness in the cockpit. *Theoretical Issues in Ergonomics Science, 12*(6), 451-471. doi: 10.1080/1463922x.2010.491874

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111*(1), 42-61. doi: 10.1037/0033-2909.111.1.42

Stanton, N. A., & Young, M. S. (1997). Validation: the best kept secret in ergonomics. Engineering Psychology and Cognitive Ergonomics, 2, 301-7.

Stanton, N. (2005). *Handbook of human factors and ergonomics methods / edited by Neville Stanton and others [eBook]*: Boca Raton, Fla. ; London : CRC, c2005.

Stanton, N., & Young, M. (1998). Is utility in the mind of the beholder? A study of ergonomics methods. *Applied Ergonomics, 29*(1), 41-54. doi: 10.1016/s0003-6870(97)00024-0

Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics, 37*, 55-79. doi: 10.1016/j.apergo.2005.06.003

Stanton, N. A., Chambers, P. R. G., & Piggott, J. (2001). Situational awareness and safety. *Safety Science, 39*, 189-204. doi: 10.1016/s0925-7535(01)00010-8

Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science, 47*(2), 227-237.

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). *Human factors methods: a practical guide for engineering and design*: Farnham: Ashgate, 2013. 2nd ed.

Stanton, N. A., Salmon, P. M., Walker, G. H., & Jenkins, D. (2009). Genotype and phenotype schemata and their role in distributed situation awareness in collaborative systems. *Theoretical Issues in Ergonomics Science, 10*(1), 43-68. doi: 10.1080/14639220802045199

Stanton, N. A., Salmon, P. M., Walker, G. H., & Jenkins, D. P. (2010). Is situation awareness all in the mind? *Theoretical Issues in Ergonomics Science, 11*(1-2), 29-40. doi: 10.1080/14639220903009938

Stanton, N. A., Stewart, R., Harris, D., Houghton, R. J., Baber, C., McMaster, R., .Green, D. (2006). Distributed situation awareness in dynamic systems:

Theoretical development and application of an ergonomics methodology. *Ergonomics, 49*(12-13), 1288-1311. doi: 10.1080/00140130600612762

Stanton, N. A., & Walker, G. H. (2011). Exploring the psychological factors involved in the Ladbroke Grove rail accident. *Accident Analysis and Prevention, 43*, 1117-1127. doi: 10.1016/j.aap.2010.12.020

Stanton, N. A., & Young, M. S. (1999). *A guide to methodology in ergonomics: designing for human use*: Taylor and Francis, 1999.

Stanton, N. A., & Young, M. S. (2003). Giving ergonomics away? The application of ergonomics methods by novices. *Applied Ergonomics, 34*, 479-490. doi: 10.1016/s0003-6870(03)00067-x

Staughton, G. C., & Storie, V. J. (1977). Methodology of an in-depth accident investigation survey (No. TRRL Lab Rept. 762).

Stelling-Konczak, A., Aarts, L., Duivenvoorden, K., & Goldenbeld, C. (2011). Supporting drivers in forming correct expectations about transitions between rural road categories. *Accident Analysis & Prevention, 43*(1), 101-111. doi: 10.1016/j.aap.2010.07.017

Sterman, J. D. (1989). Misperceptions of feedback in dynamic decision making. *Organizational behaviour and human decision processes, 43*(3), 301-335.

Stewart, R., Stanton, N., Harris, D., Baber, C., Salmon, P., Mock, M., Kay, A. (2008). Distributed situation awareness in an Airborne Warning and Control System: application of novel ergonomics methodology. *Cognition, Technology & Work, 10*(3), 221.

SWOV. (2012). Naturalistic Driving: Observing Everyday Driving Behaviour (Vol. SWOV Factsheet, pp. 5). Leidschendam, The Netherlands: SWOV: The Institute for Road Safety.

Taylor, R. (1990). Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 17 p(SEE N 90-28972 23-53).*

Taylor, R., & Selcon, S. (1994). Situation in mind: Theory, application and measurement of situational awareness. *Situational awareness in complex settings*, 69-78.

Theeuwes, J. (1998). Self-explaining roads: subjective categorisation of road environments. *Vision in vehicles--VI.*

Theeuwes, J., & Godthelp, H. (1995). Self-Explaining Roads. *Safety Science, 19*(2-3), 217-225.

Thierry, B., Béatrice, B.-A., Pierre, M., & Aurélie, B. (2009). A theoretical and methodological framework for studying and modelling drivers' mental representations. *Safety Science, 47*, 1205-1221. doi: 10.1016/j.ssci.2009.03.014

Traffic Accidents. How Safe is your Street. (n.d.). *Edinburgh Car Accidents.* Retrieved July 20, 2014, from http://accidentdatabase.co.uk/accidents/Edinburgh

Treat, J. R., Tumbas, N., McDonald, S., Shinar, D., Hume, R., Mayer, R., Castellan, N. (1979). Tri-level study of the causes of traffic accidents: final report. Executive summary.

Uchida, N., Kawakoshi, M., Tagawa, T., & Mochida, T. (2010). An investigation of factors contributing to major crash types in Japan based on naturalistic driving data. *IATSS research, 34*(1), 22-30.

Uhlarik, J., & Comerford, D. A. (2002). *A review of situation awareness literature relevant to pilot surveillance functions [electronic resource] : final report* Washington, DC : U.S. Dept. of Transportation, Federal Aviation Administration, Office of Aerospace Medicine ; Springfield, VA : National Technical Information Service, [distributor, 2002].

Uttl, B. (2005). Measurement of individual differences: lessons from memory assessment in research and clinical practice. *Psychological Science, 16*(6), 460-467. doi: 10.1111/j.0956-7976.2005.01557.x

Van, E. P., & Katel, F. (2007). Analysing 'human functional failures' in road accidents. In Y. Page (Ed.), *Traffic Accident Causation in Europe (TRACE)* (2 ed., Vol. D5.1, pp. 39). France.

Van Schagen, I., & Sagberg, F. (2012). The Potential Benefits of Naturalistic Driving for Road Safety Research: Theoretical and Empirical Considerations and Challenges for the Future. *Procedia - Social and Behavioral Sciences, 48*(0), 692-701. doi: http://dx.doi.org/10.1016/j.sbspro.2012.06.1047

Van Vliet, P., & Schermers, G. (2000). Sustainable Safety: A new approach for road safety in the Netherlands. *Traffic Research Centre, Ministry of Transport, Rotterdam*.

Vanderbilt, T. (2008). *Traffic: Why We Drive the Way We Do (and what it says about Us)*: Vintage Books.

Vidulich, M. A. (1992). Measuring situation awareness. In Proceedings of the Human Factors Society 36th Annual Meeting (pp. 40-41).

Vidulich, M., Dominguez, C., Vogel, E., & McMillan, G. (1994). *Situation awareness: Papers and annotated bibliography* (No. AL/CF-TR-1994-0085). Armstrong Lab Wright-Patterson AFB OH Crew Systems Directorate.

Von Bertalanffy, L. (1950). The Theory of Open Systems in Physics and Biology. *Science, 111*(2872), 23.

Li, W. C., Harris, D., & Chen, A. (2007). Eastern Minds in Western Cockpits: Meta-analysis of human factors in mishaps from three nations. *Aviation, space, and environmental medicine*, *78*(4), 420-425.

Waag, W. L., & Houck, M. R. (1994). Tools for assessing situational awareness in an operational fighter environment. *Aviation, Space, and Environmental Medicine, 65*(5, Sect 2, Suppl), A13-A19.

Walker, G. H. (2002). *User-Centred Car Design and the Role of Feedback in Driving.* (Doctoral Dissertation), Brunel University, Brunel. Retrieved from http://bura.brunel.ac.uk/handle/2438/6273

Walker, G. H. (2005). Verbal protocol analysis. In Stanton, N. A., Hedge, A., Brookhuis, K., Salas, E., & Hendrick, H. W. (Eds.). (2004). *Handbook of human factors and ergonomics methods* (pp. 301-3011): CRC Press.

Walker, G. H., & Chowdhury, I. (2012). Situation awareness in road design *Advances in Human Aspects of Road and Rail Transportation* (pp. 490-499): CRC Press.

Walker, G. H., Stanton, N. A., & Chowdhury, I. (2013). Self -Explaining Roads and situation awareness. *Safety Science*, 18. doi: 10.1016/j.ssci.2012.06.018

Walker, G. H., Stanton, N. A., Kazi, T. A., Salmon, P. M., & Jenkins, D. P. (2009). Does advanced driver training improve situational awareness? *Applied Ergonomics, 40*(4), 678-687. doi: 10.1016/j.apergo.2008.06.002

Walker, G. H., Stanton, N. A., Kazi, T. A., Salmon, P. M., & Jenkins, D. P. (2009). Does advanced driver training improve situational awareness? *Applied Ergonomics,40*(4), 678-687. doi: http://dx.doi.org/10.1016/j.apergo.2008.06.002

Walker, G. H., Stanton, N. A., & Salmon, P. M. (2011). Cognitive compatibility of motorcyclists and car drivers. *Accident Analysis & Prevention, 43*(3), 878-888. doi: http://dx.doi.org/10.1016/j.aap.2010.11.008

Walker, G. H., Stanton, N. A., & Salmon, P. M. (2015). *Human Factors in automotive Transport Engineering and Technology*. Farham: Ashgate.

Walker, G. H., Stanton, N. A., Salmon, P. M., & Jenkins, D. P. (2008). A review of sociotechnical systems theory: A classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science, 9*(6), 479-499. doi: 10.1080/14639220701635470

Walker, G. H., Stanton, N. A., Salmon, P. M., Jenkins, D. P., Rafferty, L., & Ladva, D. (2010). Same or different? Generalising from novices to experts in military command and control studies. *International Journal of Industrial Ergonomics, 40*(5), 473-483. doi: http://dx.doi.org/10.1016/j.ergon.2010.04.003

Walker, G. H., Stanton, N. A., & Young, M. S. (2008). Feedback and driver situation awareness (SA): A comparison of SA measures and contexts. *Transportation Research Part F: Traffic Psychology and Behaviour, 11*(4), 282-299. doi: 10.1016/j.trf.2008.01.003

Wallis, G., Chatziastros, A., Tresilian, J., & Tomasevic, N. (2007). The role of visual and nonvisual feedback in a vehicle steering task. *Journal of Experimental Psychology: Human Perception and Performance, 33*(5), 1127-1144. doi: 10.1037/0096-1523.33.5.1127

Watson, D. (2003). Investigating the construct validity of the dissociative taxon: stability analyses of normal and pathological dissociation. *Journal of Abnormal Psychology, 112*(2), 298.

Watts, B. D. (2004). *Clausewitzian friction and future war [electronic resource] / Barry D. Watts*: Washington, D.C : Institute for National Strategic Studies, National Defense University, 2004. Rev. ed.

Weber, R. P. (Ed.). (1990). *Basic content analysis* (No. 49). Sage.

Wellens, A. R. (1989). Effects of telecommunication media upon information sharing and team performance: Some theoretical and empirical observations (pp. 13-19).

Wellens, A. R. (1993). Group situation awareness and distributed decision making: From military to civilian applications. In N. J. Castellan, Jr. (Ed.), *Individual*

and group decision making: Current issues. (pp. 267-291). Hillsdale, NJ
England: Lawrence Erlbaum Associates, Inc.

Wellens, A. R., & Ergener, D. (1988). The C.I.T.I.E.S. game: A computer-based
situation assessment task for studying distributed decision making. *Simulation &
Games, 19*(3), 304-327. doi: 10.1177/0037550088193004

Weller, G., & Dietze, M. (2010). Definition, Comparison and Evaluation of Existing
Self-explaining Road Approaches in Europe *Evaluation to Realise a common
Approach to Self-explaining European Roads*: SWOV.

Weller, G., Schlag, B., Friedel, T., & Rammin, C. (2008). Behaviourally relevant road
categorisation: A step towards self-explaining rural roads. *Accident Analysis &
Prevention, 40*(4), 1581-1588. doi: 10.1016/j.aap.2008.04.009

West, R., Elander, J., & French, D. (1992). Decision making, personality and driving
style as correlates of individual accident risk. *TRL Contractor Report* (CR 309).

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple
measures. *Journal of Personality and Social Psychology*, *84*(3), 608.

Wickens, C.D. (2004). *An introduction to human factors engineering / by Christopher
D. Wickens and others*: Prentice Hall, 2004. 2nd ed.

Wickens, C.D, & Hollands, J.G. (2000). *Engineering psychology and human
performance / by Christopher D. Wickens and Justin G. Hollands*: Prentice Hall,
2000.3rd ed.

Wickens, C. D. (1992). *Engineering psychology and human performance*:
HarperCollins Publishers.

Wickens, C. D. (1995). Situation awareness: impact of automation and display
technology. *Situation Awareness: Limitations and Enhancements in the Aviation
Environment, Neuilly Sur Seine*. France.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*, 449.

Wilson, J. R., & Corlett, N. E. (2005). *Evaluation of human work / edited by John R. Wilson and Nigel E. Corlett*: Taylor & Francis, 2005. 3rd ed.

Woods, D. D., Johannesen, L. J., Cook, R. I., & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight* (No. CSERIAC-SOAR-94-01). DAYTON UNIV RESEARCH INST (URDI) OH.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. New York, NY US: Guilford Press.

Young, K. L., Salmon, P. M., & Cornelissen, M. (2013). Missing links? The effects of distraction on driver situation awareness. *Safety Science*. doi: 10.1016/j.ssci.2012.11.004

Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Hum Factors, 44*(3), 365-375.

Yu, X., Lau, E., Vicente, K. J., & Carter, M. W. (2002). Toward theory-driven, quantitative performance measurement in ergonomics science: the abstraction hierarchy as a framework for data analysis. *Theoretical Issues in Ergonomics Science, 3*(2), 124-142.

Yu, X., Lau, E., Vicente, K. J., & Carter, M. W. (2002). Toward theory-driven, quantitative performance measurement in ergonomics science: the abstraction hierarchy as a framework for data analysis. *Theoretical Issues in Ergonomics Science, 3*(2), 124-142. doi: 10.1080/14639220210124102

Zhang, Y., & Luximon, A. (2005). Subjective mental workload measures. *Ergonomia, 27*(3).