UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND THE ENVIRONMENT

The Effects of Increased Workload on Driving Performance and Visual Behaviour

by

Yan Yang

Thesis for the degree of Doctor of Philosophy

December 2011

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND THE ENVIRONMENT

Doctor of Philosophy

THE EFFECTS OF INCREASED WORKLOAD ON DRIVING PERFORMANCE AND VISUAL BEHAVIOUR

by Yan Yang

The workload of drivers has been increasing in modern times due to the growing use of In-Vehicle systems. The higher task demand resulting from these extra visual and auditory stimuli presents an increasingly challenging problem for drivers, and has become a growing safety concern, as the higher workloads can adversely affect driving performance and must be balanced against the benefits from in-vehicle systems. Existing models suggest that when the induced workload is relatively low, drivers can deal with this increased demand by using different coping strategies; however, when the workload increases above a certain level, drivers' performance decreases. These models are relatively simplistic and do not describe the extent to which drivers' coping strategies can impact on the relationship between workload and performance, which are difficult to observe by traditional measures. The literature suggests that visual behaviour or eye movements, with its physiological nature, combines attributes of both attention state and human behaviour, and can be used to provide sensitive, diagnostic, and instantaneous measurements to investigate the impact of increased workload on performance, and explore associated coping strategies.

An on-road experiment was therefore conducted to observe drivers' behaviour when their workload was increased by in-vehicle secondary tasks, and the impact this had on their performance, using eye movement as well as traditional vehicle control and manoeuvring measurements. The field experiment was run under two driving scenarios of Car-Following and Free-Driving, on two road sections in Hampshire over a period of three months, using the Transportation Research Group's Instrumented Vehicle (IV) and a FaceLab™ eye monitoring system. An Operation Simulation System was developed for drivers to perform a series of in-vehicle auditory and visual tasks through touch screen and audio systems, which reflected two different types of workload (i.e. mental and visual), with three levels of difficulty. Surveys were also conducted during and after each test run to assess drivers' workload perception and gain an understanding of their experiences of performing the tasks, and a database established to organise all the information collected to enable subsequent analyses to be conducted readily.

The results show that drivers' behaviour was significantly impacted by additional tasks, and their *secondary* task performance decreased steadily with task complexity. The effects were consistent across the two Scenarios, although driving performance generally deteriorated more for the visual tasks than auditory ones, which reflects the higher conflict of visual and manual resource caused by these tasks, and all drivers took action to compensate either by increasing their headways in Car-Following, or reducing their speed in Free-Driving. The effects were reflected in their visual behaviour, which showed higher blink rates and shrunk visual searching range for the auditory tasks, i.e. a higher mental workload over baseline driving, and higher saccade and more visual transactions between different objects for the visual ones. Differences were also found in the driving and visual behaviour of individual driver characteristics groups, including gender and different experience groups. While traditional performance measurements showed many differences in behaviour due to the extra in-vehicle tasks, the different coping strategies adopted by drivers were typically observed only through the analysis of their visual behaviour. The use of these additional measurements provides an improvement to existing models for describing the relationship between workload and performance in dual-tasking.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DECLARATION OF AUTHORSHIP

I, **Yan Yang**, declare that the thesis entitled:

The Effects of Increased Workload on Driving Performance and Visual Behaviour, and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission.

Signed: …………………………………………………………………………..

Date:……………………………………………………………………………….

ACKNOWLEDGEMENTS

I would like to thank everyone who has helped and inspired me during my doctoral study. First of all I would like to express my deepest gratitude to my supervisor, Prof. Mike McDonald, whose encouragement, guidance and support throughout this study has enabled me to develop a broad understanding of the subject. It has been an honour to work with him, and I could not have imagined having a better supervisor, whose leadership has been inspirational. I appreciate all his contributions towards my Ph.D experience, which has helped me to develop as a more-rounded academic and, more importantly, a better person. I would also like to show my appreciation to Dr. Pengjun Zheng for his ideas, time and initial encouragement, and the valuable knowledge he contributed to the experiment design and data processing.

I am extremely grateful to Mr. Doug Robinson for his enormous help in the equipment setup and his valuable comments - his knowledge contributed both to an effective design and the smooth running of the experiment. I would also like to express my thanks to Mr. Brian Mould for assembling the Instrumented Vehicle and the FaceLAB™ System. In addition, it has been a pleasure and I would like to thank all my colleagues in TRG (the Transportation Research Group), especially Dr. Nick Hounsell and Prof. John Preston, and Dr. Tim Felstead - you have all supported me when the project became tough at times, and I really appreciate that. I would also like to thank Dr. King for the encouragement, confidence in me, sound advice, time and effort in supporting me throughout my research, and to the School of Civil Engineering and the Environment at the University of Southampton, in funding me for the first 3 years of my Ph.D.

My deepest thanks also go to Dr. Kevin O'Shea for taking the time to help me with proof reading the thesis. (With the oversight of my main supervisor, editorial advice has been sought. However, no changes of intellectual content were made as a result of this advice.) Thank you also to Prof. Tiejun Cui from the Southeast University for appreciating my work, and taking an interest in our future cooperation; and to Prof. John McBride and Prof. Robert Nicholls for your constructive advice on potential research application, future work and helping me seeing the bigger picture. Thanks to Mr. Hamish Jamson and Dr. Natasha Merate from the University of Leeds for providing the individual sounds clips which were used for the design of the auditory task in this experiment. Thank you also to my external examiner, Prof. Phil Blythe for remembering me and providing valuable feedback in the final stage of this research. I would also like to thank all the subjects who participated in the study, and this PhD research would not have been possible without all of you. Finally, I would like to express my most heartfelt thanks to the Yang family, especially my parents, for all their unconditional love, support, and understanding that they gave me during my study, as well as throughout my life.

DEFINITIONS AND ABBREVIATIONS

ADAS	Advanced Driver Assistance System
AT	Auditory Task
CF	Car-Following
CrP	Correct Percentage
BAC	Blood Alcohol Content
DP	Driving Performance
DVLA	Driver and Vehicle Licensing Agency
ErD	Distance of Errors
ErP	Percentage of Errors
FD	Free-Driving
HDD	Head-Down Display
HMI	Human Machine Interface
HUD	Head-Up Display
IV	Instrumented Vehicle
IVIS	In-Vehicle Information System
LRT	Limited Resources Theory
LVT	Leading Vehicle Type
MinHW	Minimum Distance Headway
MinTHW	Minimum Time Headway
MNHW	Mean Headway
MNSP	Mean Speed
MNTHW	Mean Time Headway
MRT	Multiple Resources Theory
WMW	Whitney-Mann-Wilcoxon Test
RR	Reversal Rate
RR1.5	Reversal Rate larger than 1.5 degree
RR3	Reversal Rate larger than 3 degree
RT	Reaction Time
SDHW	Standard Deviation of Headway
SDLP	Standard Deviation of Lane Position
SDRT	Standard Deviation of Reaction Time
SDSP	Standard Deviation of Speed

SDTHW	Standard Deviation of Time Headway
SE	Steering Entropy
SP	Self-Perception
SR	Subjective Rating on workload
STP	Secondary Task Performance
TRG	Transportation Research Group
TVT	Tailing Vehicle Type
VT	Visual Task

Chapter 1 Introduction

1.1 Background

The overload of information to drivers has been a topic for traffic safety research since the late 1920's (Kutila et al., 2007), when the first radio receivers were introduced to cars. In modern times, mobile phone usage has increased steadily, and the extra workload and distraction caused by these devices whilst driving have caused wide scale research, which has led to a ban on the use of hand-held telephones for drivers in many countries (Matthews et al., 2003). For example (Rakauskas et al., 2004), research in Canada suggested that while using a mobile phone, the risk of getting involved in a serious accident increased by 38%, and drivers who used such a device were 16% more likely to cause an accident. In a similar UK study (Burns et al., 2002), drivers' performance while using a mobile phone was shown to be worse than even under the influence of 80mg/100ml of alcohol, the legal limit for the UK.

More recently, the emergence of In-Vehicle Information Systems (IVISs) and Advanced Driver Assistant Systems (ADASs) have helped drivers to reduce their travel times, provide support in navigation and for traffic information, and in communications and hazard warning assistance. However, increasing use of these systems also potentially becomes a major source of driver distraction and inattention, similar to mobile phones. They may cause an even greater impact than mobile phones due to the workload from the more intensive information they provide, especially in higher visual demands, and more complicated driver actions required. A higher workload is one of the *major sources* of inattention, which according to the National Highway Traffic Safety Administration, contributes to 25% of all police-reported accidents (Ranney et al., 2000; Wang et al., 1996) in the U.S, and another study suggested that inattention was a contributing factor to 78% of all crashes and 65% of near-crashes observed in natural driving (Dingus et al., 2006). This cause for concern is supported by similar comments from the Department for Transport in the U.K (Department for Transport et al., 2009), and it has been estimated that in-vehicle sources of inattention such as talking or interacting with instruments contribute to 55% of accidents (Wierwille & Tijerina, 1996). Figure 1.1 shows the driver cockpit in a modern vehicle, which suggests the problem of higher workload and distractions from new in-vehicle technologies may become even more critical in future vehicles, which are increasingly equipped with electronic devices such as GPS navigation systems, satellite radios and other on-board entertainment systems, MP3 players, personal digital assistants.

Figure 1.1 The cockpit in a modern vehicle, which includes ADAS and IVIS

Natural driving is already a complex task, containing high mental and visual workload elements, which requires drivers to process large amounts of information, to adapt to constantly changing environments, and be swift in decision making and responding, for example in adjusting the vehicle's speed or to manoeuvre. The workload of drivers has increased even higher in modern times, due to many compounding factors, including e.g. growing traffic density and congestion, extra road-side distractions from advertisement and Variable Message Signs, as well as mounting use of in-vehicle systems (IVISs and ADASs) - see Figure 1.2.

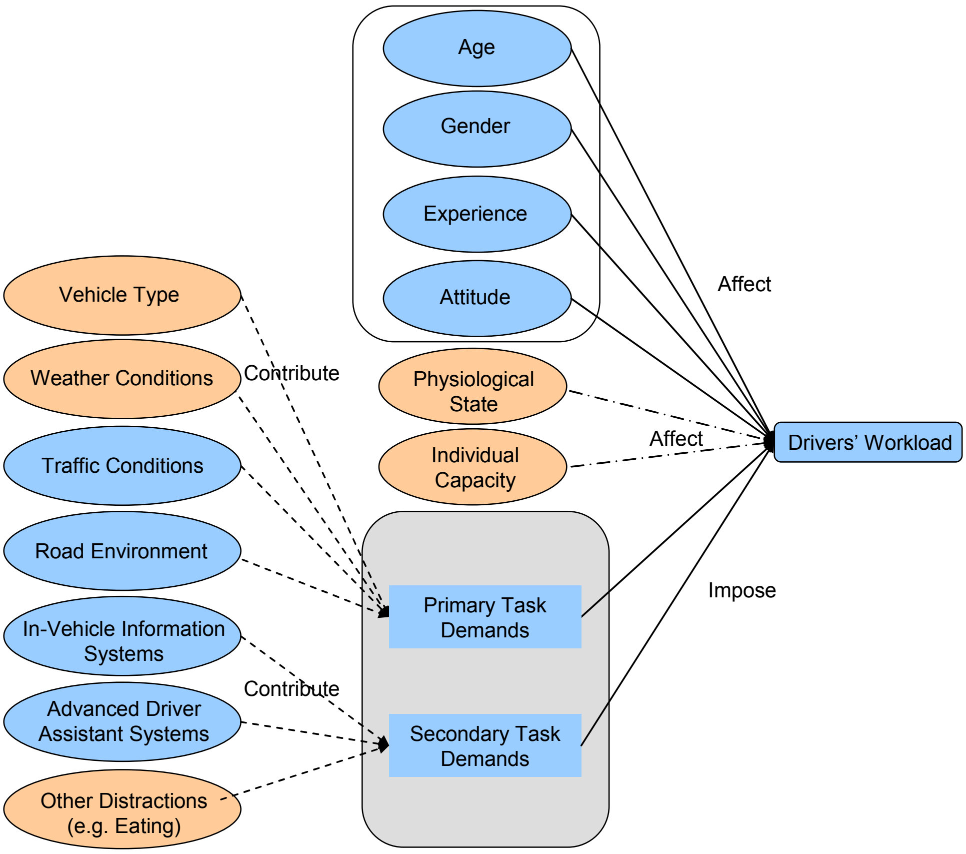


Figure 1.2 Factors which impact drivers' workloads and tasks demands

From Figure 1.2, the higher task demands caused by extra visual and auditory distractions present an increasingly challenging problem for drivers, because they impose further workload. The *task demands* (shown in the grey box in Figure 1.2) may be divided into two types: **primary** task demands, which are associated with “normal” or natural driving (i.e. vehicle control and manoeuvring), and **secondary** task demands, which result from other causes (e.g. manipulating an in-vehicle system). These task demands may interact, and are not mutually exclusive (Blanco et al., 2006). A driver’s workload is also affected by other **characteristics**, including age, gender, personal experience, driving style, physiological state (e.g. the degree of fatigue) and individual *capacity*, which is defined as the maximum limit of mental, visual, auditory and manual processing capability, and can vary from individual to individual (Wickens, 1992). (See Section 2.1 for a more detailed description of the concepts behind workload.)

An increase in drivers’ workload may inversely affect their driving performance (Meister, 1976) - see Figure 1.3 for a simple illustration. Multi- or **dual-tasking**, i.e. performing primary *and* secondary tasks at the same time, increases the drivers’ workload even further, and considerable research has been conducted to investigate how increased workloads can impact on driving performance and associated behaviour (see Section 2.1.4 for further details).

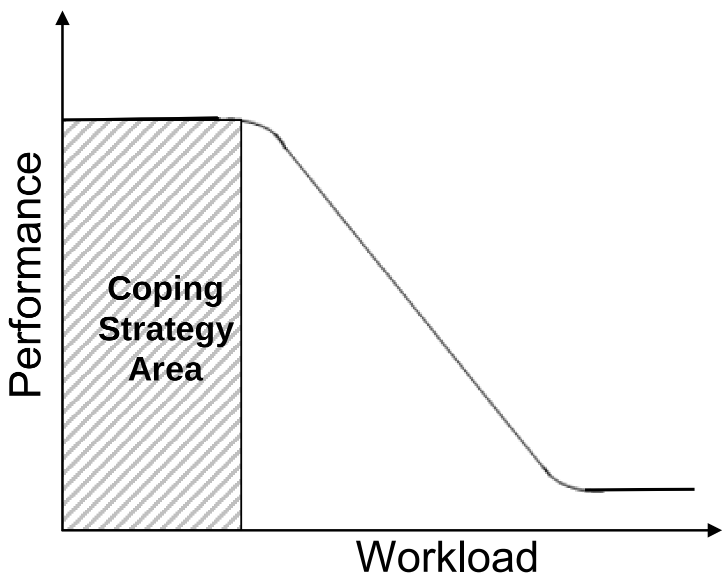


Figure 1.3 The relationship between workload and performance in driving

To cope with the increased workload, drivers adopt many different strategies to maintain a certain level of performance - see Figure 1.3. For example, they may increase their effort, to compensate for their driving performance e.g. by slowing down, or they may consciously (or subconsciously) limit the intake of superficial information, such as the condition of the road surface, or set a lower task goal e.g. a less accurate or prompt response. However, even with more effort invested, when the workload increases beyond the driver's capacity (Meister, 1976), his/her performance inevitably deteriorates.

When driving performance deteriorates, the risk of an accident due to the higher workload and other contributory causes (e.g. sudden adjacent vehicle movements) will increase - see Figure 1.4. Therefore the trend of increasing workload, particularly from the use of in-vehicle systems, is of growing concern to drivers, policy makers and researchers.

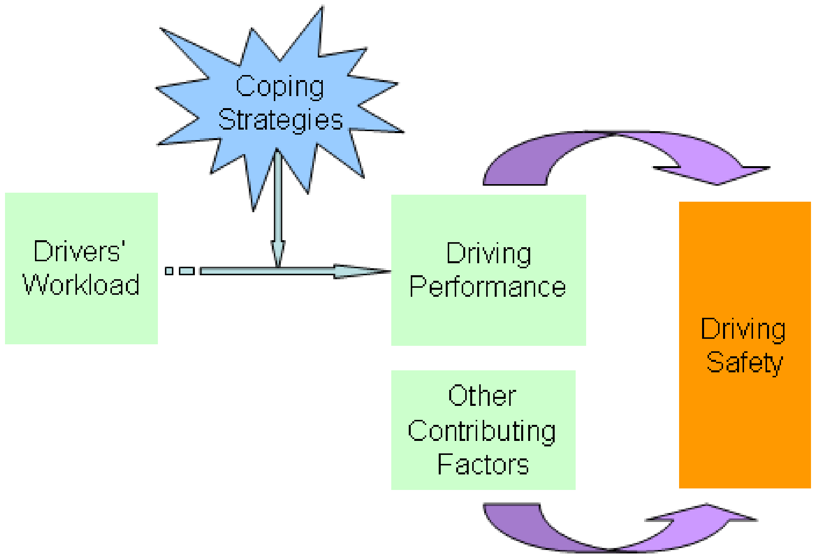


Figure 1.4 The relationship between workload, coping strategies, performance and safety

1.2 Statement of Problem

In order to understand the potential impacts of increased workload from in-vehicle systems on safety (see Figure 1.4), it is important to understand the relationships between workload and driving performance, and the coping strategies of drivers. Traditionally, such relationships have been investigated using direct measurements of vehicle control and driving performance (e.g. based on speed, headway, lateral control, and event detection). However historically, these traditional measures alone cannot provide a full explanation of the varying nature of these relationships, for example the variety of coping strategies adopted by different drivers. The relationship between workload and performance is not straightforward, and further measures are required to provide a more detailed understanding between the two.

1.3 Approach

Previous research on "eye movement" (Boersema et al., 1988; de Waard, 1996a; Lamble et al., 1999; Land & Lee, 1994; Van Orden et al., 2000) has shown the benefit of this additional physiological measurement, which provides rich information on people' attentive status (i.e. it can be used to assess workload), and is a precursor signal to how people organize and perform their activities, i.e. it can also be used as a performance measurement. **Eye movement** may be said to encompass the conscious and involuntary movement of the eyes, as well as all associated **visual behaviour**, which changes according to context, e.g. a particular distraction or the complexity of the environment. Hence the study of eye-movement offers an extra mechanism with which to investigate (and therefore provides an improved understanding of) the relationship between increased driving workload and performance - see Section 2.2 for more details.

The richer information on workload and driving performance gained from eye movement also provides typical indicators of the coping strategies used by drivers in response - see Figure 1.5. The dashed line in this figure, representing the relationship between drivers' eye movement measurements and driving performance, is one of the areas which is explored through this research.

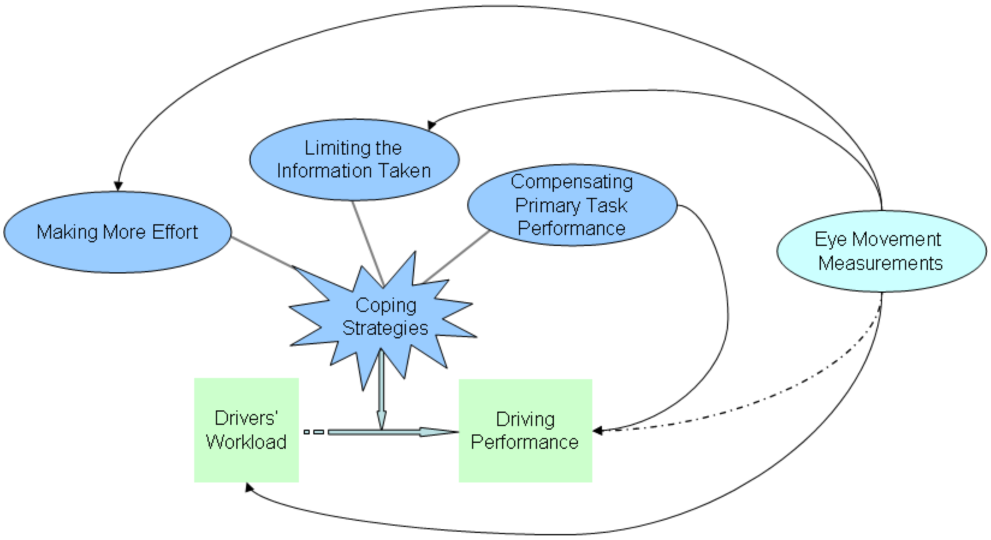


Figure 1.5 Using eye movement measurements to investigate the relationship between workload and driving performance, and drivers' resultant coping strategies

The motivation of this research is therefore to develop an improved understanding of the relationship between increasing workload and driving performance, by an analysis of drivers' eye movements, and traditional performance measures. However, it should be noted that visual behaviour patterns are task specific, and the impact of increasing workload on drivers' eye movement cannot be interpreted in isolation, but must be analysed in the context of the particular task(s) being performed, and how people react to the task (Angell et al., 2006). This research will therefore focus on the increased workload due to example in-vehicle systems whilst driving.

Driving performance also varies from individual to individual (Aarts & van Schagen, 2006a), depending on, for example age (Kua et al., 2007), gender (Brabyn et al., 2005; Dejoy, 1991) and driving experience (Chapman et al., 1998; Crundall et al., 1999a; Huestegge et al., 2010). For example, it is known that driving performance typical deteriorates after a certain age (Anstey et al., 2005). It is expected therefore that these *driver characteristics* will also have different effects on the relationship between the extra workload caused by in-vehicle secondary tasks and drivers' performance. These characteristics could also impact on the coping strategies individuals apply in order to deal with the extra workload and the *dual-tasking* situation, i.e. performing secondary in-vehicle tasks whilst driving.

1.4 Potential Benefits

This study provides an improved understanding of how the increased workload due to in-vehicle secondary tasks impacts on driving performance, and how different drivers' cope with the extra challenge in response. The results of this research will therefore help to identify the constraints and practical implications of using in-vehicle systems, which will improve their customisation and design, and help reduce their distraction and impact on driving, thereby enabling drivers to benefit from their increasing use without compromising safety - which has been an important and difficult challenge to date. For transport researchers, this study should also show the validity and benefit of using eye movement as a complementary tool (to traditional driving measurements) for investigating the response of drivers to primary and secondary tasks generally, and the impact these tasks may have on their performance. For car and device manufacturers, this study will also show the value of developing eye movement measurements as a non-intrusive process/product for helping drivers to monitor and adapt their driving, when faced with distractions and for improving safety. The results and knowledge gained through this research will therefore benefit car manufacturers, in-vehicle system developers, transport researchers, road safety advisors and, ultimately, drivers.

1.5 Scope and Objectives

1.5.1 Aim

The aim of this research is to use eye-movement as well as traditional measures:

- to study the relationship between increased workload and driving performance, by investigating the effect of in-vehicle secondary task demands, and some of driver characteristics which impact on this relationship;
- to develop an improved understanding of this workload-performance relationship, along with drivers' coping strategies and eye movement.

1.5.2 Objectives of the Research

The objectives of this research are:

- to investigate the impact of distinct types of in-vehicle secondary tasks on drivers' workload and primary and secondary task performance;
- to develop an understanding of the effects of different driver characteristics when performing these secondary tasks on performance;
- to determine how the type of secondary tasks can affect drivers' eye movement, and whether this varies according to different driver characteristics;
- to understand the relationship between drivers' performance when engaged in the secondary tasks and their self-perception of any behaviour changes;
- to assess the effectiveness of using eye movement as a tool for understanding the relationship between increased workload and performance, and drivers' coping strategies to deal with the extra workload; and
- to identify the implications and benefits of this study for different user groups.

1.5.3 Scope

This study focus on measuring the effects of in-vehicle secondary tasks on drivers' workload, and attempt to establish the relationship (and associated coping strategies) between these tasks and driving performance, by analysing drivers' eye movement and traditional measurements. The research focuses on the interaction between individual drivers and specific in-vehicle auditory and visual tasks (which *simulate* IVIS and ADAS), and the effects of different driver characteristics on the relationship between increased workload and performance. The driver characteristics considered are: gender, driving experience, age, annual mileage, and experience of and attitude towards using in-vehicle systems. For practical reasons, the impact of individual (i.e. non-task demand) factors, such as capacity and drivers' physiological state, are not considered, to limit the number of variables. Similarly, the variability due to different types of vehicles, and other secondary tasks (e.g. non-IVIS/ADAS tasks like eating or drinking) are also not considered. The factors considered or included in this study are shown in *blue* circles in Figure 1.2 further above, while those excluded are shown in orange. The effects of road and environmental factors, such as traffic density and road geometry, are not considered to be material for the purpose of this study, but have to be taken into account as contributing factors of workload, are therefore neutralised or controlled in the experimental design. It is assumed that the primary task demand of driving will also generally remain constant in one driving scenario. For example, experimental controls will be used to schedule drivers' tasks in constant driving scenarios, which involve relatively stable traffic conditions and consistent road sections.

The data collection is also conducted during the day, in dry conditions, to avoid complication with the additional workload associated with night-time and weather-related driving.

1.6 Thesis Structure

This thesis is divided into eight further Chapters:

- Chapter 2 reviews the existing literature, including why drivers' workload increases, the factors which contribute to workload (e.g. IVIS and ADAS), and the effect of increased workload on eye movements and driving performance, and how these have been measured historically;
- Chapter 3 describes the methodology and design for this research, including how the aim and objectives are addressed, and the on-road experiment and reasoning;
- Chapter 4 describes the process of data validation and reduction, including filtering, smoothing, information extraction of the raw data and the structure of the database which was used for further analysis;

Chapters 5, 6, 7 and 8 comprise the major *results* chapters, in which the analysis and findings from the on-road experiment are presented in detail:

- Chapter 5 details the findings of drivers' *subjective* workload rating and their *secondary* task performance, and the effects of different drivers characteristic on these measurements;
- Chapter 6 discusses the effect of the extra *workload* from secondary tasks on *driving* behaviour and performance, and how this relationship fluctuates according to different drivers characteristics and task difficulty levels;
- Chapter 7 presents the findings of drivers' *eye movement* while performing the secondary tasks, and explains how this relates to their attention shifting, information processing, and alertness to surroundings; it also describes how an improved understanding of the relationship between workload and drivers' performance can be obtained by eye movement;
- Chapter 8 reports the results from drivers' Post-Driving Interviews and their *self-perception* of any behaviour changes while performing the secondary tasks, which also provides further discussion of the results from the previous Chapters; and,
- Chapter 9 *summarises* the major findings from this study, including how eye movement measurements can provide an insight into the relationship between workload and driving performance, how this is of benefit to driving behaviour study, and suggests further areas for future research.

Chapter 2 Literature Review

This literature review consists of five sections, the first of which provides a review and definition of *workload*, and the existing models that describe the *relationship* between workload and *performance*, and between driving behaviour and performance. This includes a review of drivers' coping strategies and their impact on the relationship between workload and driving performance. The second section describes the existing research on *eye movement*, how this relates to workload and behaviour, and shows the potential of using these measurements to provide a better understanding of drivers' workload and performance. This includes the physical characteristics of eye movement, how it is measured, and its application to visual and driving behaviour research. Key findings relating to the impact of in-vehicle secondary tasks on driving performance are presented in the third section, while the fourth section introduces and discusses the potential research methods for this study. The final section for this Chapter then provides a summary of these discussion and the conclusions for the literature review.

2.1 Workload and Performance

2.1.1 What is Workload?

The term **Workload** is derived from cognitive and physiological theories in Psychology. It has many definitions, due to different theories and sources. For example, Parasuraman & Hancock (Parasuraman & Hancock, 2001a) suggested that "… *it is important to continue to question the very concept of workload. After many years and many hundreds of empirical investigations, we still do not have a satisfactory, consensual definition of workload* ". They view workload not as being static, or determined solely by the tasks imposed on a human "operator", but as an entity determined dynamically by many factors, e.g. those initiated by a human operator, the task in hand, the work environment, or a combination of these factors. Workload needs to be understood in the context of the **Limited Resources Theory** and the **Multiple Resources Theory**, which state that an operator's *resources* in completing ongoing tasks are limited, and human beings have the ability to activate and allocate these resources as needed. (Further explanations of the Limited Resources Theory and Multiple Resources Theory are given in Section 2.1.1.1 and 2.1.1.2 below.)

2.1.1.1 Limited Resources Theory (**LRT**)

Resources are defined (Norman & Bobrow, 1975; Wickens, 1992) as the mental and/or physical **effort** supplied by an operator to process a given task. *Resources are believed to be* *limited, and their deployment is under an operator's voluntary control* (Kahneman, 1973; Posner, 1980; Wickens, 1984a). (Effort is explained further in Section 2.1.1.3 below.) According to the *Limited Resources Theory*, the relationship between the allocation of resources and the performance of a task is linear, until the point where all resources are invested in the task. After this point, no more resources can be invested, and the task performance remains stable, assuming the task does not become more complex. The Limited Resources Theory therefore helps to explain why the task and workload of driving can remain unaffected when a *simple* secondary task is introduced. It also explains why, once the resources limit has been reached, e.g. when performing "harder" secondary tasks, either the performance of the primary and/or secondary task decreases.

2.1.1.2 Multiple Resources Theory (**MRT**)

Multiple Resource Theory was proposed by Wickens (Wickens, 1987b), who divided resources into different "resource pools", and suggested that different resource types are used for different modalities (e.g. auditory or visual) in task performing. Ongoing tasks can be further divided into two groups: auditory and visual modalities, for which auditory and visual resources are invested respectively. In addition to auditory and visual, the *central* resource is required to perform all types of tasks. According to MRT theory, when two tasks have an overlap in terms of resource requirement, either the primary or secondary task, or both, will be affected since the resource will become fully allocated. When two tasks require different resources, for example when one is visual and the other is auditory, there will be no direct conflict of resources, unless the performance of either task is constrained by the central resource limitation -see Figure 2.1.

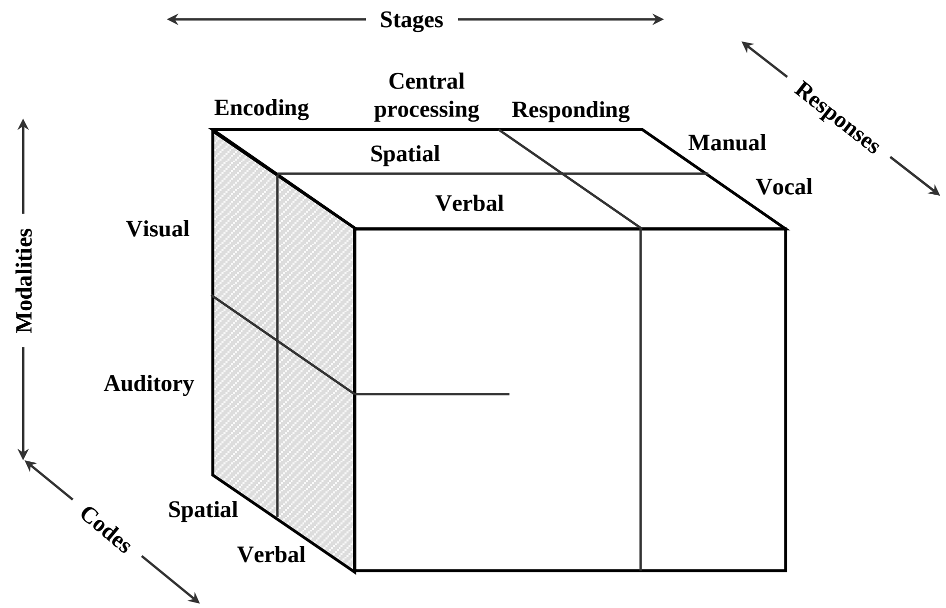


Figure 2.1 The proposed structure of processing resources *(Wickens, 1984b)*

Figure 2.1 shows how the concept of multiple resources can be represented as points in a three-dimensional space, where the axes represent the input information *modality* (e.g. visual or auditory) and associated *codes* (e.g. spatial or verbal), the *stages* of human processing (encoding, central processing, responding), and the associated *response* types (e.g. manual or vocal). The closer two tasks are in this space, the more they will interfere with each other, as they compete for the shared type of input, processing and output resources.

Results from dual-tasking in driving studies (Brooks, 1968; see also Baddeley, 1976; Wickens & Liu, 1988) strongly supported this theory by pointing out that verbal and spatial processing use different resource pools. In accordance with MRT, "*the performance decrements are less severe during the concurrent performance of cross-modal tasks e.g. visual intensive and auditory, compared to intra-modal ones, for example both tasks requiring density visual information processing*" (Wickens et al., 1983).

A modified MRT has been proposed in recent research, in which verbal and spatial information processing are considered as **working memory** types in dynamic control tasks, such as driving (Angell et al., 2006). *Working memory* refers to the temporary storage and information processing while tasks are being performed (Baddeley, 1996; Salmon et al., 1996). During driving, the *spatial* working memory is *essential* to the perception of position, speed, and the acceleration of adjacent as well as driver's own vehicle; while *verbal* working memory is required for tasks like reading road signs and to perform other in-vehicle tasks (e.g. listening to the radio).

In driving, it is crucial for drivers to use *vision* to perceive the road scene, and *spatial working memory* to judge relative vehicle position, and *manual controls* to adjust steering, accelerator and the brakes, which correspond to the three dimensions of visual information modality, spatial working memory and manual response in Multiple Resource Theory, i.e. perceptual, cognitive, and physical. Therefore, secondary tasks which fall into the same dimension as "normal" driving will be expected to have the highest overlap of resources (de Waard, 1996a), and the higher competition for these resources will occur, and cause the highest workload for drivers.

2.1.1.3 Effort

Generally, capacity is believed (Wickens, 1992) to be flexible before reaching the upper limit of resources, i.e. when processing simple tasks, even if the tasks become more demanding, the availability of resource can be increased, although this requires more **effort**. In workload-related research, the term *effort* describes the mobilisation of additional resources as a compensatory process (Aasman et al., 1987), and since it reflects the reaction to demands and is expandable (de Waard, 1996b), it is considered one of the most important components of *mental* workload. According to Mulder (Mulder, 1986), there are *two* forms of effort: *task-related* effort, or that required for information processing (e.g. computation), and *state-related* effort, or that required for individuals to adjust their energetical state (e.g. to stay alert against fatigue). Vicente et al. (Vicente et al., 1987) suggest two things should be noted about effort. First, the amount of effort expended by an operator is *not* necessarily related to the task demand, but depends more on the *nature* of the tasks, which is related to e.g. experience and the operator's state. For example, in driving specifically, different drivers may apply different amounts of effort in order to perform the same task; and tasks which are performed while driving may compel different levels of effort, even though they are similarly demanding when performed statically. Secondly, there is no simple relationship between the effort invested and the operator's performance.

Only *task-related* effort will be studied for the purpose of this research, which is not concerned with drivers' physiological states. Therefore, the term used in this thesis refers to task-related effort, which includes the *amount* of resource invested in task performance, *where* resources are deployed, and *how effective* the effort is for dealing with an increased workload. As it is task-related, the effort drivers invest in a specific task is considered part of their driving strategy, and can be detected and analysed by using either performance or eye movement measurements.

2.1.1.4 Definition of Workload in This Research

De Waard (de Waard, 1996b) suggests workload can be simply defined as a demand placed upon a human, which is *exclusively* attributed to external sources. However, Rouse et al. defined workload by its *indication*, i.e. "*workload is not only task-specific, but also person-specific"(Rouse et al., 1993)*. Hence individual operator capabilities, states, mood, etc., as well as the task complexity, have an effect on the workload experienced. This is in line with others (Kahneman, 1973; O'Donnell & Eggemeier, 1986), who defined workload as being directly related to the *proportion of the capacity* that an operator spends on performing a task. Therefore, in these definitions, workload not only depends on the task demand, but also the *amount* of resource operators invest in performing the task.

When applied to driving, de Waard defined workload as the proportion of *information processing* capacity used to perform tasks (de Waard, 1996a), and when a secondary task is performed in addition to driving (Tijerina, 1996), workload can be defined as the competition of resources between these tasks. Similarly, Angell (Angell et al., 2006) defined workload as the competition in driver resources between primary driving and other distractions, e.g. turning on the radio.

Therefore, workload is defined as the amount of driver's resources (perceptual, cognitive, and physical) that are used to perform primary and secondary tasks concurrently, and depends on the individual, as well as the task demands, and the interaction between them. The same task demands do not necessarily result in an equal level of workload for different individuals, and hence the assessment of workload must take into account drivers' characteristics and the task difficulty, i.e. the interaction between drivers and the task demands.

2.1.2 Benefits of Workload Research

In an extensive review of workload and the human operational system, Parasuraman and Hancock (Parasuraman & Hancock, 2001b) commented that there have been many reasons why we should understand mental workload. One is that a detailed knowledge of mental workload is needed to help improve existing information-processing models of human behaviour, which are insufficient. Another is the practical issues associated with human capacity in driving, that is, when more and more in-vehicle systems are introduced, driving performance can decrease as a consequence, and this has become a safety concern. The assessment of workload is therefore essential to the design of in-vehicle systems to minimise the impact on driving performance, which can lead to the concept of a maximum and a minimum workload for drivers to sustain an acceptable level of performance.

In addition, although there are currently no validated functions that relate workload measurements precisely to accidents (Angell et al., 2006; Brookhuis & deWaard, 2001). High mental workloads are believed to cause conditions including insufficient attention, imperfect perception and inadequate information processing (Brookhuis & deWaard, 2001), and this can increase the likelihood of human errors, which is one of the most significant contributors to accidents. Therefore an understanding of what causes high mental workloads, particularly when performing in-vehicle tasks, can be of benefit to road safety.

Finally, research into the measurement of workload can be used as an indicator to reduce the conflict or adjust the nature of task demand for in-vehicle systems, and this has benefits for their customisation. For example, research can be undertaken to detect and predict a driver's workload and apply this information to activate and control in-vehicle systems accordingly. Another example is the use of a "personal log-in system" which can target workload-sensitive drivers, and constrain their use of in-vehicle systems if necessary.

2.1.3 Driving Behaviour and Performance Models

Increased workload impacts on many aspects of driving behaviour. However, this research focuses only on driving performance change (e.g. vehicle control) rather than specific driver behaviour (e.g. hand movements) because, according to the Driver-Vehicle System Theory (see below), it is reasonable and practical to consider the driver and the vehicle as a whole. In addition, driving performance can be measured and assessed using several parameters in different *levels* of behaviour, as suggested by certain Driving Behaviour Models, which are described in Section 2.1.3.2 further below.

2.1.3.1 The Driver-Vehicle System Theory

In order to perform a driving task (Senders et al., 1967a), drivers need to receive input from their driving environment, interpret it, make decisions, and then carry out actions. In theory, the actual movement of a vehicle therefore depends on drivers' perception of the environment, their information processing and decision making, as well as their actions on manoeuvring (such as moving a foot from the throttle to the brake, as well as other forms of kinematic control). However, for the purpose of driving behaviour research, existing practices suggest (Zheng, 2003) it is neither practical nor possible to observe every component of a real driver-vehicle system in every way. Zheng (2003) summarised that "*the rationale of doing so has been justified in the crossover model of manual control behaviour where McRuer et al (1959) found that the human operators can adjust their performance and respond in such a way as to make the total open-loop transfer function of human-machine system behave as a first-order system and thus the behaviour of a driver-vehicle system can be better discussed under the framework of manual control".* Most existing models therefore focussed on both the driver and the vehicle as a whole unit (Cody & Gordon, 2007), and a simple driver-vehicle system model is shown in Figure 2.2.

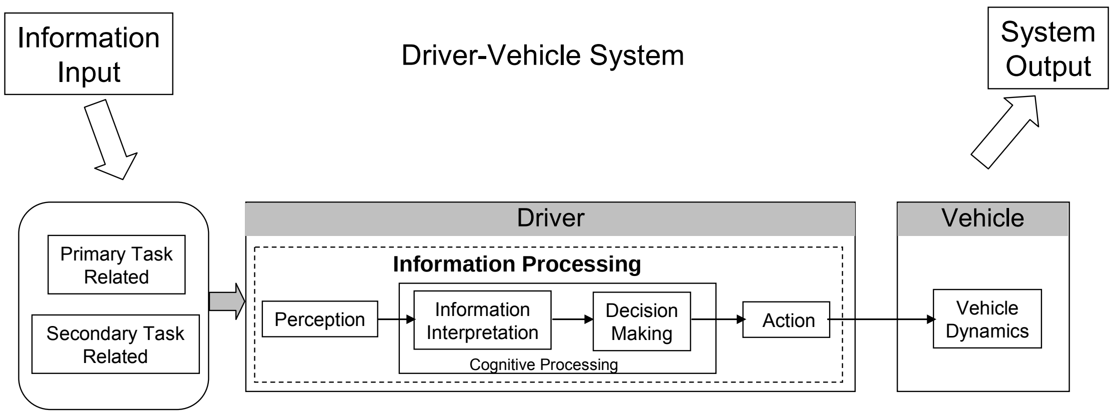


Figure 2.2 A Driver-Vehicle System Model (from Zheng, 2003)

This relatively simple and descriptive model provides the theoretical basis for generating the parameters to assess driving behaviour and performance. In the model (see Figure 2.2), the workload induced by primary and secondary tasks, as well as the interaction between them, may be considered as a *system* input, while the vehicle dynamics may be considered as the output. Therefore, drivers' performance in both tasks can be measured through e.g. the longitudinal and lateral control of the vehicle.

2.1.3.2 Driving Behaviour Models

The primary driving task is defined by Parkes as "*safe control of the vehicle within the traffic environment*" (Parkes, 1991a). A more detailed, hierarchical model of driving behaviour was provided by Michon (Michon, 1985), where driving a car was described at three levels. At the top or *strategic* level, high-level decisions are made, such as the route-choice. At the intermediate or *manoeuvring* level, reactions to local situations take place, including reactions to the behaviour of other traffic, such as changing lanes. At the lowest or *control level*, basic vehicle-control processes occur, such as lateral-position control. At the lower levels, the automation of processes may occur as experience develops, especially when a higher level control process is required. An alternative way to look at the driving performance is to use Rasmussen's three levels of performance (Rasmussen, 1987), in which the human behaviour was categorised into three hierarchies: the *knowledge-based* behaviour, the *rule-based* behaviour and the *skill-based* behaviour, corresponding to the three abovementioned levels respectively.

Michon and Rasmussen's hierarchy models therefore provide a basis for investigating drivers' performance *and* behaviour (de Waard, 1996a), as *performance* measurements can be connected to the three levels of *behaviour*. For example, steering wheel adjustments can reflect driving performance at the lowest level, i.e. the control level, while effectiveness in car-following and mirror checking can reflect performance at the manoeuvring level, and an effective route choice reflects performance at the strategic level. However, demands can exceed capacity in all three levels of behaviour, and may result in performance deterioration.

The increased workload caused by secondary tasks is most likely to affect behaviour at the manoeuvring level, i.e. the performance of "*rule-based behaviour*" or more specifically, the maintenance of a safe speed and/or headway in current road and traffic conditions (Lunenfeld & Alexander, 1990). However, according to the hierarchy models (Michon, 1985; Rasmussen, 1987), actions at a higher level can affect behaviour at a lower level. Therefore, the impact of increased workload from secondary tasks should be reflected in performance measurements at the two lower levels of behaviour, i.e. manoeuvring and control, which includes longitudinal and headway keeping and mirror checking (for manoeuvring), and lateral control and steering wheel adjustments (for control).

2.1.4 Relationships between Workload and Performance

This section provides a review on previous models describes relationships between workload and performance, and research investigating effects of higher workload on driving performance. The high workload imposed on drivers was separated into two categories according to their source: those caused by primary driving demand (e.g. complex road conditions or busy traffic), and those due to secondary task demands (e.g. turning on the radio or in-vehicle conversation). When reviewing the effects of secondary tasks, some essential concepts and theories, for example, *divided attention* and *task switching*, *time sharing* and *task paralleling* are also described.

2.1.4.1 Models of the Relationship between Workload and Performance

A simplified relationship between workload and performance was described by Meister (Meister, 1976), in which three regions were defined, region A, B and C, see Figure 2.3. Region A is described as when workload is in a relatively low level, even when it increases, performance can still stay unaffected; Region B represents the area when workload increases upon a certain threshold, performance started steadily decreases with the increasing workload; While in Region C, when workload exceed an upper limit, performance remains at this minimum level, and does not deteriorate even further.

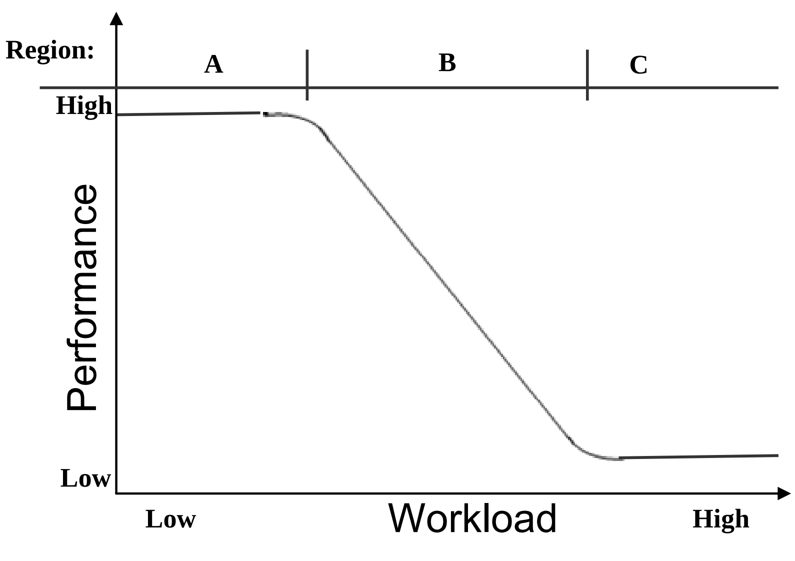


Figure 2.3 The hypothetical relationship between demand and performance

(based on Meister, 1976)

This model only describes the effect of task demand which is "*too high*" for a human operator. However, other research also suggested that not only highly demanding tasks have negative impact on task performance, but also the low demanding ones, for example, boredom can reduce the capacity by lower operators' arousal level, and in turn cause a higher mental workload (Meijman & O'Hanlon, 1984). Later, Brookhuis's (Brookhuis et al., 1999) studies of drivers' behaviour under *underload* and *overlaod* conditions also suggested that both of these conditions were related to driver performance impairment, as the underload condition causes reduced alertness and lowered attention. De waard (de Waard, 1996a) therefore included a Regions D and A1 into the previous model to form a "Inverted U" Shape, as suggested by Teigen (Teigen, 1994), shown in Figure 2.4. In driving, for example, the under-load condition (Region D) can occur when driving in a dull environment (motorway driving), or when assisted by an adaptive cruise control system, which could reduce a driver's arousal level and cause fatigue; while the overload condition (Region B and C) can occur when dual-tasking, which can also lead to diverted attention from the road ahead, or insufficient time for critical information processing. Between underload and overload (Region A1, A2 and A3), drivers can compensate for the increased workload by re-allocating effort. Therefore in these Regions, adding further task demand may increase workload without necessarily affecting driver's performance. When the task demand carries on increasing and exceeds a certain level, a driver will be unable to maintain the performance even with the maximum effort, and driving performance will deteriorate unavoidably, as shown in region B in Figure 2.4.

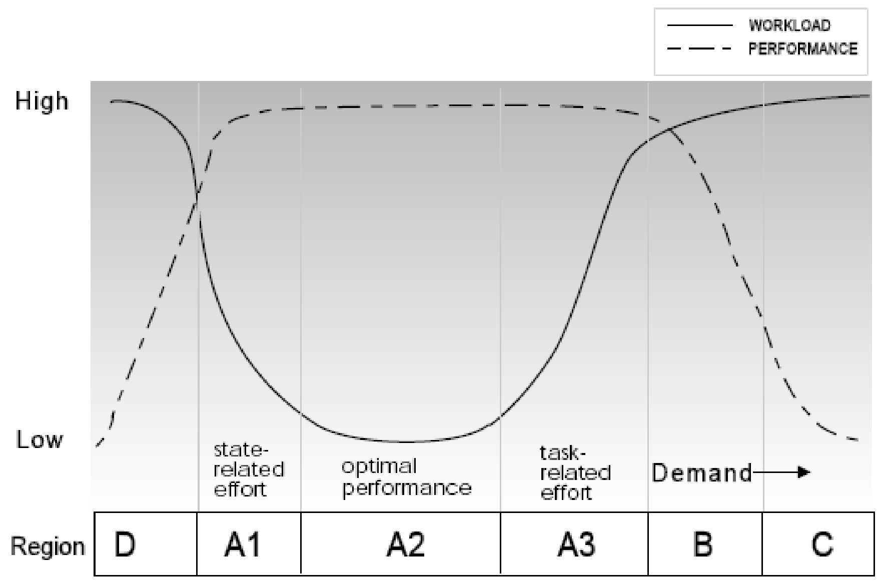


Figure 2.4 The inverted-U relationship between workload and performance *(de Waard, 1997)*

Another model described the relationship between workload and performance **-- Malleable Attentional Resource Theory** was proposed by Young and Stanton to describe the mental underload when using automated systems (Young & Stanton, 2004). Usually, automated systems perform well when a human operator is able to cope with them. However, they are often ill-prepared to cope with a sudden demand. The poor performance with automation includes the fatigue, poor situational awareness, too much trust, and vigilance decrements. The **Malleable Attentional Resource Theory** contains three basic rules: *1) the attentional resources are malleable; 2) the attentional resources are yoked to task demand; and 3) there is a lag in the attentional resource expansion.* In other words, when task demand is reduced (e.g. when assisted by an automated system), the attentional resource pool shrinks to accommodate the reduced demand (i.e. as being malleable), which is considered to be cognitively efficient by Young and Stanton; and when an increase in demand suddenly occurs to the human operator, e.g. actions required in a semi-automated system, the individual will be unable to cope with this requirement, since the resource pool cannot expand quickly enough (Young & Stanton, 2004). This theory provides an expanded explanation for the effect of underload condition in Region D and A1 in the “inverted U”.

2.1.4.2 Effect of High Workload on Performance

As discussed above, it should be noted that the workload does not always cause performance deterioration. When workload increases, driving performance may decrease, but not always, for example, as suggested in the previous section in Region A. Reid and Nygren also suggested (Reid & Nygren, 1988) that even when *an operator’s workload exceeds a certain value, it is the probability of performance breakdown increase, rather than that there is a value that will definitely result in performance breakdown.* To avoid overload, drivers may apply different strategies, like making more effort, or simplifying the task by limiting the number of information sources they use for decision making (Cooper & Zheng, 2002; Wright, 1974). However, these strategies can also result in impaired driving if a wrong decision is made due to the reduction of information source, for example, neglecting the factor of road surface condition (dry or wet) when distracted (Cooper & Zheng, 2002). High workload could lead to risky behaviour, for example, Hoyos (Hoyos, 1988) also noted that high workload can decrease drivers’ hazard detection, and increase the probability of a driver making risky decisions due to time pressure. This along with other evidences, suggest that high workload may be related directly to high accident risk, for example, Wood et al. (Wood et al., 2003) found in a simulator study, that drivers took longer to brake for critical events (e.g. stop signs or pedestrians at intersections) when they were given the task of counting the number of a certain type of pedestrian. In addition, even when sometimes extra workload appears to be not high enough to cause driver's overload or performance deterioration, the cumulative effect of sustaining attention or making effort in a higher level during certain period of time can also be a source of stress which can reduce drivers' attentional capacity (Hancock & Warm, 1989), which in turn may lead to safety concern.

2.1.4.2.1 Impact of Primary Task Demands on Performance

Workload caused by complex driving conditions, i.e. primary driving related workload, can impact on both secondary and primary task performance. For example, Brown & Poulton (Brown & Poulton, 1961a) reported that subjects' performance in calculation tasks decreased, in terms of completed fewer, in more complex driving environments than in simpler ones. Lee and Triggs (Lee & Triggs, 1976) reported that drivers missed more target lights in a peripheral detection task while driving through busier and more complex environments. The former also found a lower speed in the complex driving environments (i.e. city traffic) than in simpler ones (residential areas). Harms (Harms, 1991) also found that in more complex driving environments (e.g. different village areas and close to rural junctions), reaction time to calculation tasks was longer, in combination with a lower driving speed than in the *control* condition (or *baseline*, i.e. normal driving without performing secondary task), on highway sections. Martens and van Winsum (Martens & Van Winsum, 2000) found negative effects on both reaction time and hit rate for a peripheral detection task when the driving task difficulty was increased by external causes such as narrow curves or the appearance of an unexpected obstacle. Other research on distractions (Engström et al., 2005; Horberry et al., 2006; Jahn et al., 2005; Jamson & Merat, 2005; Tornros & Bolling, 2006) reported that increasing task demand in relatively complex situational environments will impair driving performance even further. For example, in Jamson and Merat's (Jamson & Merat, 2005) research on distraction due to surrogate IVISs tasks which was conducted in a simulator, both secondary task and primary performance deteriorated, but even more so in more complex road sections (curve or with hazardous events). Especially when performing visual tasks the deviation in lane position only increased by about 2cm compared to baseline (from 20cm to 22cm, values estimated from figure) on straight road sections; however, on curved road sections the measurement increased by 9cm, from 36cm to 45cm.

However, drivers can adapt their behaviour for the workload caused by complex driving conditions, as Senders et al. stated (Senders et al., 1967a), the attentional demand from road environment depends not only on the road condition itself (e.g. layout, geometry, etc), but also on the traffic situation which includes driver's speed choice and other factors. Therefore the speed limit at which the drivers travelled can be determined by the drivers' information processing capacity, according to the information generation rate of the road. Drivers tended to drive at the speed which can balance these two factors (Senders et al., 1967a), for example, drivers behave differently when travelling on familiar or unfamiliar routes.

2.1.4.2.2 Impact of Secondary Task Demands on Driving Performance

The effect of workload induced by the secondary task on driving performance was found in many studies. Either consciously or subconsciously, drivers developed a strategy to reduce the primary task load whilst performing the concurrent secondary task. Harbluk, Noy, Trbovich, Eizenman (Harbluk et al., 2007a) investigated the influence of the mental tasks by arithmetic tasks, and results showed that more occurrences of hard braking were found during the most difficult mental tasks, along with the effect of narrowing of attentional focus. Effects of secondary tasks on driving performance were also reported in research related to mobile phone use (Alm & Nilsson, 1994; Brookhuis et al., 1991; Burns et al., 2002; Poysti et al., 2005; Strayer et al., 2006; Summala et al., 1996; Tornros & Bolling, 2005), In-vehicle entertaining systems (Summala, Nieminen & Punto, 1996), Surrogate IVISs systems, both auditory and visual tasks (Anttila & Luoma, 2005; Engström, Johansson & Östlund, 2005; Jamson & Merat, 2005; Östlund et al., 2004), and other in-vehicle systems (Burnett & Joyner, 1997; Horberry et al., 2006; Jahn et al., 2005; Serafin et al., 2007; Walker et al., 1991; Wierwille & Tijerina, 1996; Wittmann et al., 2006; Zheng et al., 2008).

Driving a vehicle in real traffic is predominately visually demanding and, according to Wickens (Wickens, 1984a) resources are multiple, and therefore simultaneous processing of visual information for a secondary task causes a structural interference with driving. On the other hand, when tasks are presented verbally and require only vocal responses, they cause less intrusive stimuli and reduce the interference with primary tasks, which should maintain the structural interference to a minimum. However, as mental tasks still conflict with driving in mental resources, drivers may therefore attempt to free up resources by simplifying the primary task, which can be manifested by changes in driving behaviour. Performing visual tasks generally causes more behaviour changes and performance deterioration, in terms of lower speed, higher speed deviation and reversal rate, impaired lane control, delayed braking and less awareness of surroundings, so that drivers react slower or miss potential hazards.

To understand the underlying process and effects of dual-tasking, some essential concepts including attention dividing, task switching, time sharing and task paralleling, are briefly described as following:

Divided attention and task switching

Higher workload caused by secondary tasks is not only due to the task demand itself, but also to the effect of attention being divided between the IVIS systems and the primary task. Divided attention requires the ability to process and/or respond to information while conducting two tasks simultaneously, which is related to a decrease of driving capacity (Lengenfelder et al., 2002). When two tasks are performed simultaneously, the combined workload can be more than the sum of workload of the two tasks, the effect of which is more pronounced for elderly people (de Waard, 1996a). Anderson and colleagues (Anderson et al., 1998) proved the cost of divided attention by interrupting a memory task in two stages (encoding and retrieving), which suggested that dividing attention at the encoding stage disrupted memory performance, and the reaction time was prolonged in the retrieving stage. The longer reaction time again was more noticeable for older drivers. Schvaneveldt (Schvaneveldt, 1969) showed that even performance on relatively simple tasks can be degraded when they are performed simultaneously with other complex and independent tasks. Switches between tasks were also associated with persistent costs, even when the two tasks are predictable and simple (Rogers & Monsell, 1995). Given that driving already requires attention being divided, including continuously monitoring on-road information, while scanning the environment for potential hazards, and shifting attention onto dashboard as needed (Lengenfelder et al., 2002), adding another task will cause further cost for drivers. In addition, it was found that one the most pronounced effect on multi-tasking is time pressure (Moray et al., 1991), which is an essential element for road safety.

Time sharing

Edquist stated (Edquist, 2008) that to *successfully perform the primary driving task, drivers must be able to select the necessary information and process it in time to make the appropriate decision and execute the required action.* Therefore anything that interferes with this process and/or the speed of this processing, for example performing secondary tasks, will impair the driving performance. The degree of interference between concurrent tasks can be manipulated by the task structure and task priority. As described in Section 2.1.1.2, in the three-dimensional space composed of the information modality (e.g. visual/auditory), the stages of processing (encoding, central processing, responding), and the response types (e.g. manual/vocal), the closer the two tasks are, the more they will interfere to each other due to more competition for shared resources.

Tsimhoni and Green (Tsimhoni & Green, 2003) *explained that while performing visual concurrent tasks, there are four constructs that may affect the time-sharing while driving: 1) the time pressure of looking at the road, 2) the interference of concurrent driving in cognitive* *demands, 3) the drivers' ability of postponing information processing, and 4) the cost of "chunking" the task into smaller parts.* Obviously when drivers have to look at something inside the vehicle, they cannot simultaneously look outside the vehicle, and hence their ability to respond to external events suffers. In addition to the effect of visual-manual conflict, many driving simulator studies have also shown that even non-visual tasks can affect the visual perception as well, because of the mental resource competition.

Task paralleling

Previous research discovered that there is one exception of human's limitation in information processing in task sharing, which is tasks which can be processed in parallel due to atomisation developed by practice (Shiffrin & Schneider, 1977). The dynamics of visual attention also allows people to shift attention quickly within fixations, including objects in periphery vision. Schumacher suggested that (Schumacher, 2001) *skilled procedural decision making and response selection for two or more tasks can proceed at the same time under adaptive executive control*. After some training, some subjects could achieve time sharing in the dual-tasking, and also, interference between tasks could be adjusted by task priorities and personal preferences. According to Michon's *control-manoeuvre-strategy* model (Michon, 1985), the control level as the lowest level of driving behaviour is the highly likely to be performed automatically by experienced drivers. It is possible that experienced drivers can use automated vehicle control when coping with in-vehicle tasks based only on the peripheral view (Summala, Nieminen & Punto, 1996), so that the multi-task demands in normal driving conditions still remain in their attentional capacity.

However, there are limitations of these parallel controls. When driving demands are increased by, for example, traffic density increasing, driving in the intersection, traffic circle, etc., the tasks demands may sometimes exceed driver capabilities (Hancock & Parasuraman, 1992), and the useful visual field still gets narrower with increased mental workload (Jahn et al., 2005).

2.1.5 Drivers' Coping Strategies

In theory, task performance is supposed to be *linearly* related to the amount of resources allocated on performing it, until the moment that all resources are invested (Norman & Bobrow, 1975). However, this association does *not* necessarily always occur, since humans can apply different coping strategies to deal with higher workload, for example, by *effective* time sharing and task paralleling, unaffected performance could occur even when a second task was added. Drivers can also adapt their behaviour and cope with the task demands by changing their strategies as necessary. In a comprehensive review by Parasuraman and Hancock's

(Parasuraman & Hancock, 2001b), a matrix of associations and dissociations on the performance and workload was presented, which showed that the response from human operator is adaptive to task demand in nature, and the human component in any human-machine system should *not* be considered as a simple linear element.

Drivers' strategies when coping with a secondary task demand include, making more effort or adjust performance target, for example, compensating the primary task, using less demanding strategies (e.g. effective time sharing and task paralleling), postpone or cancel the secondary task processing or other subsidiary activities (Cnossen et al., 1997). When the task demand is in the range that drivers can cope by investing more effort, the performance will remain unaffected. In this case, the performance measurements will not reflect any change. However, a latent performance decrement may occur due to accumulated pressure on a driver. In practice, a method to detect this effort is required. On the other hand, even when the workload is very high, drivers may employ some compensation strategies to protect their primary driving tasks by, for example, dropping slowing down, increase safety margin, omitting event detecting or information exacting (e.g. reducing peripheral vision), as Sender and colleges (Senders et al., 1967) suggested, by applying these strategies, drivers can, to some extent, maintain the workload in an acceptable level. Previous researches (Jamson and Merat, 2005, Brookhuis et al., 1999, Rauch et al., 2008) have also found the similar strategy when coping with in-vehicle secondary tasks. In this case, similarly, some driving performance measurements can be equally insensitive. However, these strategies may not always be enough to deal with the high level workload, or drivers tend not to, or fail to apply appropriate strategies in real driving (Horrey and Lesch, 2009), therefore the performance decrease regardless (Yang et al., 2009a, Rakauskas et al., 2008). For example, Horrey's study revealed that even when they were fully informed of the on-coming high demanding road environments, drivers tended not to postpone the secondary task even when they could.

Strategies also differ between individuals, and some strategies are more effective than others therefore require less effort. Many researchers have proposed the concept models of the relationship between workload and performance; nevertheless, due to the strategies involved, the relationship itself is not straight forward. Also, as drivers' effort and strategies cannot be observed from the traditional performance measurements alone, the validation of these models needs a new method which can provide quantitative support. A more build-in measurement is needed in order to investigate the effort and coping strategies for an improved understanding of this relationship.

2.1.6 How is Workload Measured?

Because workload is a multi-dimensional phenomenon, which is related to *performance* (e.g. changes in both primary and secondary tasks), to *subjective experiences* (e.g. mental effort and time pressure), and to *physiological states* (e.g. stress and effort), measuring/assessing workload involves various methods including *performance criteria* (e.g., the quantity and quality of performance), and *subjective evaluation* (e.g., the ratings of effort level) (Schvaneveldt et al., 1998) and *physiological measurements* (e.g. heart rate and alert level). An overview of all these three categories of measurements is given in this section, with some of them being described in more detail.

2.1.6.1 Self-Report on Subjective Experience

Assessing workload based on self-reported experience is one of the most commonly used methods. According to Kahneman (Kahneman, 1973), the definition of mental workload is directly related to the *proportion* of the capacity that an operator can spend on the task in hand, and the measurement of mental workload is the *specification* of that proportion (deWaard & Brookhuis, 1997). The workload changes sometimes are not easily observed in driving performance measurements because drivers are inclined to *actively* cope with the increasing task demands by adapting their behaviour (Cnossen, Brookhuis & Meijman, 1997). However, the increased workload can be apparent in self-reports of the drivers due to the individual experience when dealing with the task, for example, on time pressure, stress or more information processing required. As de Waard suggested (de Waard, 1996b), no one is able to provide a more accurate judgement with respect to the experienced workload, especially mental workload, than the persons themselves. Since they are the closest to the essence of workload, Sheridan cited in Wickens, (Wickens, 1984b), considered self-reports as the best measure for workload, although de waard (de Waard, 1996b) suggested that most self-report measures are not sensitive in the A2 region in Figure 2.4. The commonly used self-report methods are summarised in the following sections.

2.1.6.1.1 RSME (Rating Scale Mental Effort)

RSME is developed by Zijlstra (Zijlstra, 1993), including a scale from 0 to 150 mm, with a cross line in every 10 mm, and the higher effort is indicated by higher scale. Along the line, from 1 to 120 mm, 9 levels of statements are given, from “absolute no effort” to “extreme effort”. The amount of effort invested into a task has to be indicated at one point along the line. In this scale, the effort is considered as a whole, i.e., there is no separate evaluation for different types of workload, e.g. the physical and mental workload, known as uni-dimensional scale.

2.1.6.1.2 Activation Scale

The activation scale is comparable to RSME, i.e. a uni-dimensional scale which requires subjects to mark on the line. However, instead of the effort level statements, on Activation Scale, the reference comments along the line are daily activities like "I'm reading a newspaper", and the marking for subjects should be relative workload compared to the workload from the stated activities. The scale has a range from 0 to 270 and the measures are taken by the distance from the point 0.

A similar ratings system, knows as Dual-Task Activation Rating, was proposed and used effectively in a previous research on distractions during driving (Angell et al., 2006). In the Dual-Task Activation Rating, ratings are collected immediately after each secondary task. Drivers are asked to rate the workload caused by a secondary task by comparing the experience of completing this task with a reference task (for example turning on the radio), which has a fixed value of 100. Such Activation Ratings should be an optimal method for gaining feedback from drivers of their experience due to the minimised delay in subjective rating, and it can be easily collected during driving test as ratings are only given by the driver vocally, and are recorded by the experimenter on a questionnaire-style sheet. An example of Dual-Task Activation Rating is provided in Appendix B.

2.1.6.1.3 NASA-Task Load Index (TLX)

NASA-TLX is a multi-dimensional workload assessment scale (Hart & Staveland, 1988), which has been employed in numerous studies over the past 20 years. It was developed by the Human Performance Group at NASA Ames Research Centre. NASA-TLX was originally designed for assess pilot workload in the aviation domain, and later was used largely in ergonomics research. NASA-TLX provides an overall workload score based on a weighted average of rating on six subscales: mental demands, physical demands, temporal demands, the own performance, effort, and frustration.

2.1.6.1.4 Driving Activity Load Index (DALI)

Driving Activity Load Index (DALI) is a revised vision of NASA-TLX, which was especially designed for the IVIS-related workload measure in driving circumstance, especially when assess the workload due to using in-vehicle systems (Pauzie & Pachiaudi, 1997). In DALI subjective rating scale, there is a scale rating procedure for six pre-defined factors: the effort of attention, visual demand, auditory demand, temporal demand, interference, and the situational stress. The rating on these factors is followed by a weighting procedure in order to combine the six individual subscales into a global score (Pauzie, 2008).

2.1.6.1.5 SWAT (Subjective Workload Assessment Technique)

The SWAT workload assessment is a multi-dimensional Subjective Workload Assessment Technique (Reid & Nygren, 1988). It is a subjective rating technique that uses three levels: low, medium, and high, for each of three dimensions of time load, mental effort load, and psychological stress load to assess workload. It uses conjoint measurement and scaling techniques to develop a single, global rating scale with interval properties. The use of SWAT entails three steps: the first step is called scale development. All possible combinations of three levels of each of the three dimensions are contained in 27 cards. Each operator sorts the cards into the rank order that reflects his or her perception of increasing workload. The second step is the event-scoring, that is the actual rating of workload for a given task or mission segment. In the third step, each three-dimension rating is converted into numeric scores between 0 and 100 using the interval scale developed in the first step.

2.1.6.1.6 Workload Profile (WP)

WP was developed by Tsang and Velazquez (Tsang & Velazquez, 1996) based on the Multiple Resource Theory, MRT (Wickens, 1987b). The Workload Profile asks the subjects to provide the proportion of attentional resources used after they had experienced all of the tasks. The tasks to be rated are listed in a random order down the column and eight workload dimensions are listed across the page. The workload dimensions used in this technique were defined by the resource dimensions hypothesised in the MRT: perceptual/central processing, response selection and execution, spatial, processing, verbal processing, visual processing, auditory processing, manual output, and speech output. The definition of each dimension is given to subjects when they are required to rate on the tasks. In each cell on the rating sheet, subjects provide a number between 0 and 1 to represent the proportion of attentional resources used in a particular dimension for a particular task. A rating of “0” means that the task placed no demand on the dimension being rated; a rating of “1” means that the task required maximum attention. The ratings on the individual dimensions are later summed for each task to provide an overall workload rating.

2.1.6.1.7 Findings from Self-Reporting

Generally, mental workload assessed by self-report reflect the task demand level, where ratings in control conditions (i.e. with out secondary tasks) are lower than that in dual-tasking (Lansdown, Brook-Carter, and Kersloot, 2002; Gopher, 1990), lower for simpler tasks than for more complex tasks (Lee et al., 2001), and lower for less resource-conflicting tasks (e.g. auditory tasks) than for high conflicting ones (e.g. visual tasks). In addition, greater time pressure was reported to be significant while undertaking the dual task conditions as opposed to the primary task alone (Lansdown, Brook-Carter, and Kersloot, 2002). The results from drivers' self-reported performance changes also suggest that drivers are able to recognise their impaired driving performance when conducting secondary tasks (Anttila & Luoma, 2005; Blanco et al., 2006; Jamson & Merat, 2005; Ma & Kaber, 2005). However, there is also evidence from self-reports to suggest that drivers tend to underestimate the potential severity of these distractions and continue to multi-task (Anttila & Luoma, 2005; Blanco et al., 2006; Liu, 2001).

2.1.6.2 Task Performance Measures

2.1.6.2.1 Secondary Task Performance Measurements

Secondary task performance can be used to indicate workload level because human beings only have a limited amount of information processing capacity, and after drivers have managed to perform the primary task, the "unused" capacity or *Spare capacity* (Brown & Poulton, 1961b) is largely available for secondary tasks. Therefore the performance of secondary tasks will reflect changes in the spare capacity and total workload. When interacting with the in-vehicle systems, driving performance is likely to degrade, as the driver devotes greater attention to use the device and less to the other task, the secondary task performance will decrease. The commonly used measurements are: secondary task completion, accuracy (i.e. percentage of correct responses or errors), reaction time, etc. Measurements of secondary task performance have shown to be sensitive to task complexity, task type, and interface design. For example: Some drivers voluntarily skipping secondary tasks or having the experimenter request to skip them because of safety concerns when performing complicated tasks (Blanco et al., 2006). Drivers shown to be more accurate when responding to message content for messages presented in the synthetic speech (concatenate) compared to the recorded human voice (Harbluk & Lalande, 2005), suggesting one caused less workload than the other. Anttila and Luoma 's (Anttila & Luoma, 2005) found in a field study that, both in visual (arrow) and in auditory continuous memory (cognitive) task, the percentage of correct responses was the highest for the static situation and the lowest when conducted in the urban environment comparing to rural and motorway. The percentage of correct responses is decreased with the increasing task difficulty. These results suggested that the secondary task performance could be an effective measurement for drivers' workload.

2.1.6.2.2 Peripheral Detection Task (PDT) Measurement

PDT is another commonly used method for workload evaluating from the secondary task performance measure family. It has been confirmed to be valid as an objective measurement for workload or "spare capacity". It uses various visual secondary tasks (e.g. flashing lights requiring drivers' responses) as indicators of concurrent visual demand. Since the driving task requires continuous visual information processing, this method reflects the spare visual capacity, and was therefore applied for the safety evaluation of IVISs. Apart from the visual demand itself, some research also found that PDT is also sensitive to cognitive load, as it was found that mental tasks can affect drivers' eye movement patterns and peripheral vision (Miura, 1986). Crundall and colleagues showed in various studies (Crundall, Underwood & Chapman, 1999a; Crundall et al., 1999b) that drivers' PDT task performance decreases when there were traffic hazards presented on videotapes.

However, PDT tasks sometimes still show the biased sensitivity of visual workload over mental or other sense modalities. It was found by Harms and Pattern (Harms & Patter, 2003) that only navigation messages including visual demand, i.e. visual navigation messages and combined visual and verbal messages, had a statistically significant effect on PDT performance. Thus, it is questionable to use PDT-task as the only standard measures for the workload increasing. Apparently, a method which is less dependent on one specific mode (visual) would be more reliable and more general for measuring workload. In addition, the extra visual distraction caused by the PDT task limited the application of PDT in the real-road driving condition.

2.1.6.3 Physiological Measures

Sometimes, when workload change is not easily shown in either secondary task or driving performance, since drivers can actively cope with the task demands by adapting their driving behaviour, they can be apparent in physiological measures (Cnossen, Brookhuis & Meijman, 1997), for example, heart rate or eye movement. Different physiological measures have been found to be differentially sensitive to either a global level or a specific stage in information processing. One of the advantages of physiological measures is that they do not require the drivers' response, also they can be collected continuously, and most of them are not invasive to task performing. Historically, the workload can be measured by the Cardiac Functions, Background Electroencephalogram (EEG), Blood Pressure, Respiration, Skin Resistance Response and Skin Conduction Response, Hormone levels, and Electromyogram, etc. (O'Donnell & Eggemeier, 1986). In this section, only the eye movement measure is described.

The eye movement can actually be considered as both performance and physiological measures (de Waard, 1996b). Many researches have used eye activity measures that correlated with cognitive demands to measure the real-time workload (Ahlstrom & Friedman-Berg, 2005; Van Orden, Jung & Makeig, 2000; van Orden et al., 2001; Brookings, Wilson, & Swain, 1996; Wilson, 2001; Wilson & Caldwell, 2002). Findings indicate that the *blink rate*, *blink duration*, and *saccade duration* (saccade: rapid eye movement) decrease while the *pupil diameter*, number of saccades, and frequency of *long fixations* are increased with the increased visual workload (Iqbal et al., 2005; Rognin et al., 2004; Van Orden, Jung & Makeig, 2000; Veltman & Gaillard, 1998; Zeghal et al., 2002). However, for some eye movement measurements, the mental workload may have the opposite effects as that in the visual workload (Recarte & Nunes, 2003; Recarte et al., 2008). There is a need for careful explanation of the eye movements in tasks as complex as driving, where multidimensional workload is intertwined. Same as other physiological measures, eye movement measurements are constitutions and not invasive, which provide the opportunity of assess workload in real-time. A brief overview of some eye movement parameters can be used for workload evaluation is presented below, and the visual functions, measurements, and correlation with workload and performance are discussed in more detail in Section 2.2.

2.1.6.3.1 Fixation Measurements

Fixation is the maintaining of the visual gaze on a single location. Eyes typically alternate between saccades and visual fixations. Therefore, fixation is normally defined as the pauses intervals between discrete jumps of eye movement (i.e. saccades) during a specific visual task. It is assumed that during these pauses the eyes are to be relatively stable, i.e., “fixated” (Inhoff et al., 2008), and it is during fixation when the visual information is being extracted. Fixation duration and frequency are two commonly used parameters. In researches of observing eye movement patterns (Wilson & Eggemeier, 1991), the frequency of fixation was found to be related to the importance of viewing of instruments (i.e. the more important information being viewed more often), while the length of fixation was generally believed to be related to the difficulty of information extraction (i.e. the more difficult it is, the longer it takes to extract useful information).

2.1.6.3.2 Pupil Diameter Measurement

Pupil diameter increases as a result of *Sympathetic Nervous System* (SNS) -- *innervated* muscle groups cause pupil dilation. Kahneman (Kahneman, 1973) suggested that the increased task demands and increased resource investment were reflected in the increases of pupil diameter. The same relationship between workload and pupil diameter was also found by other researchers (Recarte et al., 2008). However, the largest changes in pupil diameter can be caused by other reasons, e.g. changes in illumination. It is therefore suggested that the pupil diameter is suitable only in laboratory situations (Kramer, 1991).

2.1.6.3.3 Eye Blink Measurements

Blink is *endogenous* eye movement behaviour. The blink rate and blink duration are studied as sensitive measure for workload changes (Recarte et al., 2008; Velichkovsky et al., 2002). Hancock used the eye-blink frequency as the secondary task measures and found that turn sequences were associated with greater demands on central attentional capacity than straight driving (Hancock et al., 1990). Eye movement measures used to be considered only useful in the assessment of visual demands (Kramer, 1991), in recent years, some researchers have provided evidence that blink can also be a sensitive measure to auditory or cognitive demand situations (Recarte et al., 2008; Yang et al., 2009).

2.1.6.3.4 Eye Saccade Measurements

Saccade is rapid eye movement. Larger saccade amplitude is related to higher visual searching demand. Huestegge (Huestegge et al., 2010) found that saccade amplitude is significantly higher when viewing images with high hazard than those with medium hazard. A recent research found that the correlation between peak of saccade velocity and subjective mental workload during risky behaviour (Di Stasi et al., 2009).

2.1.7 Summary of Workload/Performance Review

In summary, it is known that driving performance generally deteriorates with an increased workload due to primary and secondary task demands. There exists a need to define the criteria of driving performance which may lead to accidents, when the impairment is below a certain threshold (Brookhuis et al., 1999). However, due to the difficulty in workload assessment and drivers’ coping strategies for dealing with increased workload, in view of the large amount of literature, an understanding of the relationship between the increased workload and driving performance which includes a quantified effect of drivers’ coping strategies and driver characteristics is still unavailable. In other words, it is still unclear how, and to what extent, increases in workload contribute to the decrements in driving performance, and how drivers react to the increased workload to maintain their driving performance. Pervious research has explored the use of visual or eye movement parameters that correlate with both cognitive demands. Hence there is the potential for using eye movements, along with other measures to evaluate drivers’ workload, and also to investigate drivers’ coping strategies.

2.2 Eye Movement

As discussed in the previous section, eye movement is proposed to be the more sensitive, built-in, and instantaneous measure required as a real time link between workload and driving performance. From the literature review, the physiological nature of eye movement enables the capability of measuring both. More importantly, previous research provides evidence of it being developed to understand coping strategies.

Eye movement provides information on the *vision distribution* to organise a given activity (e.g. in what sequence a specific task is completed *and* how much information is needed in each step to perform this task), which is not easily obtained through other methods. It was believed that "*observing operators' eye movement can be used as a natural tool to understand how mind acquires and processes visual information*" (Shinar & Gurion, 2008). In addition, visual data can reflect the moment-to-moment information processing (Rayner, 1998), and is closely linked to real time workload (Shinoda et al., 2001). Therefore, the analysis of eye movement can reveal attention-allocation, which is the key to understanding drivers' coping strategies.

To state the problem clearly, several questions need to be answered:

- What is eye movement?
- Why do we study eye movement -how is eye movement related to workload and behaviour?
- How is eye movement connected to driving performance?
- How do we measure eye movement?

In the following sections, 2.2.1 describes the physical characteristics of eye movement, and the above questions are answered by literature reviews on eye movement research, in Sections 2.2.2, 2.2.3, 2.2.4, and 2.2.5 respectively. A summary of these reviews is provided in Section 2.2.6.

2.2.1 Physical Characteristics of Eye Movement

This subsection explains what eye movement is by exploring its physical characteristics.

2.2.1.1 "Saccade and Fixation" Strategy

One of a widely accepted theory is that there is a consistent pattern of eye movement throughout the animal kingdom which can be refer to as "saccade and fixate" strategy (Land & Horwood, 1995; Land, 2006). **Saccade** is the fast eye movements (has a velocity normally as high as 500° per second), which direct eyes from one point to another; while **fixation** is the intervals between saccades, in which gaze is held almost stationary, normally with a duration for about *200-300ms*. However, in experimental observations, even in the duration of fixation, eyes are never perfectly steady. Therefore, fixation is defined as the viewing intervals between discrete jumps of eye movement (saccades) during a specific visual task. It is therefore assumed that during fixation the eyes are to be immobile, i.e. “fixated” (Inhoff et al., 2008). It is during fixations that information is taken in, and during saccades we are effectively blind because the speed is too high for the eyes to extract any information. Therefore, fixation is a parameter for us to study how visual information is extracted, while saccade is more a parameter of how drivers organise their searching behaviour.

More details about saccade can be found in an intensive review on eye movement research (Rayner, 1998). *Saccades need to be distinguished from three other types of eye movements:* ***pursuit, vergence,*** *and* ***vestibular eye movements****. Pursuit eye movements occur when our eyes follow a moving target, therefore with markedly slower velocity than saccades, however, if the target is moving quickly across our visual field, we often make saccades to catch up with the target (White, 1976). Vergence eye movements occur when we move our eyes inward, toward each other, in order to fixate on a nearby object. Vestibular eye movements occur when the eyes rotate to compensate for head and body movements in order to maintain the same direction of vision. Although pursuit, vergence, and vestibular eye movements are important and extensively studied, saccadic eye movements are more relevant in the typical information processing tasks.* When the visual information requires more searching, saccade occurs more frequently, and therefore, the less efficient the visual information processing becomes. When a secondary task requires visual information processing, drivers need to make more saccades to perform the searching action required. However on the other hand, because of the low efficiency during saccade (i.e. we do not obtain any visual information, and even before and after, visions are blurred), when the visual information process is intensive, the so called saccade suppression may occur, where saccade is inhibited (Riggs et al., 1974). Two thresholds of saccade in different researches areas were applied: in research in reading, for example, a threshold of gazes move larger than 2 degrees is generally used as a definition of saccade, i.e. eye movements which are larger than 2° is defined as saccade (Rayner, 1998). In the research of dynamic activities (e.g. driving), because of the constant eye movement during these activities, depending on the task condition, the criteria based on fixed gaze angle is not always sufficient. Saccade is therefore defined either according to the range of the gaze angles, or based on the velocity of the eye movement, for example, movement faster than 500° per second is defined as saccade (Carpenter, 1998).

2.2.1.2 Saccade Durations and Fixation Duration

There is a debate around whether saccade durations should be included in the computation of gaze duration or not. Most researchers followed the lead of Just and Carpenter (Just & Carpenter, 1980) and used only fixation duration values for **gaze** durations, and the transition process is excluded in the definition of fixation duration. In this research, the same definition will be applied, i.e. fixation is considered the same concept as gaze, which refers to the activity of eyes fixed on a specific object, the processes of moving to that objects and leaving away from it are not counted as a part of fixation or gaze.

2.2.1.3 Fovea, Parafovea and Peripheral Vision

Fovea, parafovea and **peripheral** are three regions of the visual field. *Fovea* is the vision centre, with an angular diameter of between 0.31° and 2°. Fovea is the most acute visual region, even though it is only about 1/4000th of the retinal surface (Steinman, 2003). Away from the foveal centre is the *parafovea*, where the acuteness starts falling rapidly compared to fovea. Parafovea extends out to 5° on either side of fixation, and then *peripheral* indicates the region beyond the parafovea, where the vision is even poor. Therefore, to see details in a specific object, one needs to move the fovea to the target of interest. During driving, drivers need to move their eyes from point to point to use the fovea vision effectively. Less crucial information, for example monitoring adjacent vehicle speed, can also be extracted by peripheral vision, especially for experienced drivers (Underwood et al., 2002).

Research on physical nature on eye movement suggested (Land, 2006) that there were similarities in terms of the mechanisms that control eye movement in all daily behaviour. In tennis playing, for example, it was found that for a good player, eyes always move to the bouncing point one second before the ball arrives at that point (Rayner, 1998); while similarly, when making tea (Land, 1999) and a sandwich (Hayhoe, 2000), eyes are always moved prior to the behaviour for the next step, by moving gaze from one task-relevant object to the next, ignoring all other objects that were not involved in the activities. These findings revealed how study on eye movement can provide the information of how activities are organised, how information is visually collected and reflect the strategy of dealing with an ongoing task, and therefore, how eye movement study can be used in workload and behaviour/performance research.

2.2.2 How is Eye Movement related to Workload and Behaviour

2.2.2.1 The Strategic Nature of Eye Movements

Eye movement reveals the behaviour strategy in reading, picture viewing and piano playing (Land, 2006), where *top-down* behaviour patterns were found. It was only after portable eye movement monitoring equipment was invented in recent years that eye movement has been studied in daily activities like walking, tea and sandwich making, and in sport. A considerable amount of research have suggested that eye movement having the ability to reveal detailed behaviour strategy when organising activities and performing tasks. Yarbus (Yarbus, 1967) investigated visual searching patterns when looking at pictures with a number of different questions in mind. These questions were related to different aspect of the image, for example the people, or the clothes they wear. The results showed that each question evoked a different pattern of eye movements, related to the information required to answer the question. This meant that eye movements are not simply related to the structure of the picture itself, but also to top-down instructions from the brain. Similarly, drivers also view the scenes with different questions, for example: is it safe to change lanes? Yarbus' work proved that eye movements were reflex actions, but much more strategic in character. Such a strategy of eye movement determines that it is close related to workload and behaviour.

2.2.2.2 Eye Movements, Workload, and Behaviour

The previous study shows that eye movements can be both the workload and performance measures, for example, Recarte (Recarte et al., 2008) has taken blink rate as a reliable and nonintrusive workload measurement; while in some researches, eye movement was used as part of the task performance measurement (Angell et al., 2006). The physiological nature of eye movement provides the opportunity of building up a bridge between drivers' workload and their performance in a detailed level.

Although it was argued that it is not always necessary to move the eyes to identify every object and human beings can move attention without necessarily moving their eyes (Posner, 1980), it is still believed by researchers that with complex stimuli, it is more efficient to move eyes than to move attention (He & Kowler, 1992; Rayner, 1998). In complex information processing tasks such as driving, where the link between the two is significant, analysing eye movement behaviour can provide the information on attention allocation. "*Even though there is limited systematic study of where drivers look in traffic, from observations it is clear that drivers visually detect the places from which they need to obtain information, e.g. on the car in front, the in-vehicle display, the pedestrians and cyclists, the road signs and traffic lights, etc*." (Land, 2006). How this information is taken, especially when dual-tasking and how drivers cope with the extra task demand can be revealed by observing drivers' eye movement.

2.2.2.3 Eye Movements in Driving

Eye movement research started about 20 years ago. Huey (Huey, 1908) cited in (Rayner, 1998) first observed the role of eye movements in reading. Later the research interest spread out into reading in playing music, daily activities, sports, etc. Findings from these researches revealed that more complicated visual content typically caused more and longer fixations. Since 1990s, research on the effects of cognitive load (e.g. solving math and physics problems) on eye movement lead to the concept that eye movement can reflect not only visual, but also mental workload. It is only since head-mounted eye trackers became available in the 1980s that it has been possible to study active tasks such as driving (Land, 2006).

Driving a vehicle in a real traffic situation is predominately visually demanding and, according to Wickens (Wickens, 1984a), the resources required in driving were multiple and the concurrent processing of visual sources would cause structural interference. Even though the commonly believed statement that 90% of the information required for driving is visual was questioned by a review of vision in driving (Sivak, 1996), it is still true that most driving-critical information is collected visually. The driving task involves dealing with the road itself (steering, speed control), other road users (vehicles, cyclists, moving and stationary pedestrians), paying attention to road signs and other relevant sources of information. It is thus a varied task, and it is expected that a range of eye movement strategies will be employed (Land & Horwood, 1995; Land, 1992; Mourant & Rockwell, 1970; Rockwell, 1988).

Early studies on eye movement were mainly conducted on road sections with low curvatures (e.g. typically US roads), therefore only a weak correlation between gaze direction and steering was found (Zwahlen et al., 2003). The relationship between steering and gaze was later clarified by findings from sharp curves demanding more attention (Land & Lee, 1994). Land and Lee (Land & Lee, 1994) conducted their research on a very winding road, and the results revealed that the recorded gaze direction and steering wheel angle were very similar, as shown in Figure 2.5. In particular, drivers spent much of their time looking at the “tangent point” on the approaching bend.

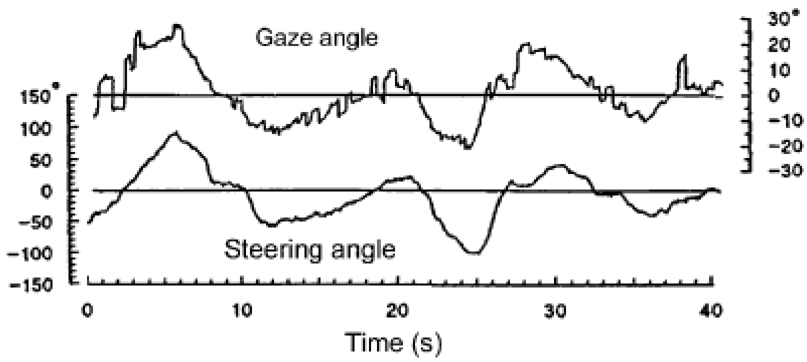


Figure 2.5 Gaze angle on winding road *(Land & Lee, 1994)*

In normal driving, drivers spend most of their time looking ahead with about 5° down from the horizontal, in vertical gaze angle. It was reported (Hughes & Cole, 1986) that when watching a movie film of driving view, 25% of fixations were located in the central region, and 80% of the remaining were within 6°of the visual centre, i.e. only 15% of fixations were beyond 6°. It is believed that looking away from the road ahead for any period of time is detrimental, as Summala has found (Summala, 1998) the lane keeping on a straight road deteriorated as a function of drivers fixating away from the view ahead.

To keep a good steering, drivers need the information of further vision: *tangent point*, as well as near road region: *road edge* (Land & Horwood, 1995); when only the further region was visible but the near region was not, curvature matching was still accurate but the lane control was poor; On the other hand, with only the near road region visible but without the further region, lane maintenance was good, but curvature matching was poor, mainly due to large and abrupt steering induced by the short time available for reaction to the road edges (<0.5 s). Therefore neither the further vision, nor the near road lane edge input is sufficient on their own, but the combination of the two allows fluent accurate driving. The result suggested that about 5° down from the horizon gives a good result on both criteria, explaining why 5° down from the horizon was found in normal driving.

Peripheral vision can be developed to monitor lane position (Mourant & Rockwell, 1970). Learner drivers first use foveal vision for lane keeping, then increasingly move foveal gaze to more distant road regions, and learn to use their peripheral vision to stay in lane. Summala et al. (Summala, Nieminen & Punto, 1996) reached similar conclusions. This may also be related to the fact that experienced drivers show a much larger searching area (Crundall, Underwood & Chapman, 1999a).

2.2.2.4 The Effects of Secondary Tasks on Eye Movement

Typically, the percent of concentration on the road centre, searching area and fixation on the in-vehicle display are used to measure the demands of secondary tasks (Victor et al., 2005). When visual tasks became more difficult, drivers looked less at the road centre area ahead, and looked at the display more often and for longer durations with higher deviation; while for the auditory task, gaze concentration on the road centre area increased with the increase of task complexity. A similar effect of gaze concentration due to mental demand (as measured by a visual perimeter) were previously reported by Rantanen and Goldberg (Rantanen & Goldberg, 1999). As a consequence of shrunken gaze area, mental tasks can also cause decreases of visual detection. Olsson and Burns (Olsson & Burns, 2000) found that counting backwards interfered with the detection of peripheral lights. Strayer, Drews and Johnston (Strayer et al., 2003) also reported that when participants were involved in a hands-free phone conversation, they responded slower to the leading vehicle's brake lights. Horberry and colleagues (Horberry et al., 2006) related to the visual impairment caused by mental workload directly with road accidents. They found in a study conducted in a simulator that engaging simulated hands-free mobile phone conversation impaired drivers' responses to pedestrians crossing the road.

More evidence of the linking of eye movement parameters and performing secondary task were proposed by other research. For mental tasks, Recarte and Nunes (Recarte & Nunes, 2000; Recarte & Nunes, 2003) found that the fixations were longer and saccades were smaller for participants. Drivers were found to check mirrors more often (Recarte & Nunes, 2000), and their saccade rate decreased (Harbluk et al., 2007b). As for visual tasks, Harbluk, Noy and Eizenman (Harbluk et al., 2002) found that additional tasks reduced the overall number of saccades, scanning to the periphery, and checking instruments and mirrors, and that these changes were related to an increase in hard braking. In one of the researches carried out in an aviation simulator by Federal Aviation Administration, USA (Ahlstrom & Friedman-Berg, 2005), it was found that when detecting a higher visual condensed task, operators' blink duration and pupil diameter increased, but no significant changes were found in blink frequency, saccade frequency and saccade distance.

In summary, a review on the effect of secondary tasks on eye movement suggests that secondary tasks increased driver workload, decreased the area of visual space, and caused drivers to perceive less of visual objects. It should be noted that different types of tasks have very different effects on eye movement. In order to use eye movement measurements to investigate the workload and performance, tasks which cause the visual behaviour change need to be categorised and the characters of various eye movement parameters in each type of task have to be studied carefully.

2.2.2.5 The Effects of Different Driving Situations on Eye Movement

Apart from the effects due to secondary tasks, eye movements are also affected by different driving situations. For example, when driving at a lower speed (e.g. 30 mph or less), eyes are much more freed up for the multiple demands of dealing with other road users and potential obstacles, than driving at a higher speed (Land, 2006). Because at low speed, steering only requires peripheral lane edge information, and the need for monitoring distant tangent points is much reduced. Therefore, driving at a lower speed (normally accompanied by less workload) is observed to produce a wider visual searching area.

Visual searching behaviour in driving is similar to any other daily activity, showing the "top-down" strategy of behaviour. Drivers' visual fixation patterns were found to be affected by various demands in a systematic and predictable fashion (Land & Horwood, 1995; Recarte & Nunes, 2000; Velichkovsky et al., 2002; Victor, Harbluk & Engstrom, 2005). Searching for required information (e.g. road signs) is also influenced by the complexity and familiarity of the environment. In a more complex visual context, targeting and extracting the useful information can be more demanding. As the demands of traffic situations increase, peripheral vision is sacrificed in order to provide greater attention for information uptake by the fovea (Miura, 1986). This is confirmed by Shinoda et al. (Shinoda, Hayhoe & Shrivastava, 2001), who found that the detection of stop signs requires an active search, and the visibility of the signs depends on the concurrent visual background. Active searching also depends on what information is needed by the driver, for example, on a familiar route the need for checking road signs may be not as much as on an unfamiliar one. As Shinar and Gurion has found (Shinar & Gurion, 2008), drivers' fixation were *less dispersed* on the above and right of the roadway (where the traffic signs are located in USA), and were more concentrated further down the drive lane when they became more familiar with the route. During driving, time pressure can also affect drivers' visual searching strategies in different driving conditions. In free-driving, vision focus is on a location further down the road; however, when following other vehicles, drivers tend to concentrate on the rear of the leading vehicles (Shinar & Gurion, 2008). The visual drift was explained as the drivers' *initiative strategy of obtaining information with maximal lead time*.

Previous research showed the sensitivity of visual behaviour on road environment, which implied that eye movements cannot be solely investigated, but needs to be considered under a specific driving situation. For any on-road research, to obtain high quality eye movement data, the road environment need to be controlled carefully.

2.2.3 Eye Movement Measurements

Eye movements were monitored in many different ways. The apparatus of eye tracking can be divided into three types according to the techniques they rely on: (a) the *surface electrodes* placed around the eyes, which are good at measuring saccade behaviour but accuracy of the gaze location measuring is deficient; (b) the *contact lenses* attached to the surface of eyes; (c) the *video-based* non-contact, corneal reflection method, in which, infrared light is reflected from the eyes and sensed by a video camera or some other specially designed optical sensors. The third one is widely used in researches because it is non-invasive and inexpensive. The FaceLab^{TM} system used in this research is based on the third video-based method. The parameters of eye movement which were used in workload and performance related researches have been reviewed intensively in this section. Although there has been research concerning the measurements, evaluation, and reporting of eye movement data, no standard for eye movement measurements have been adopted. Therefore, some of the most commonly used measurements in secondary task evaluation are introduced below. The ones already described in Section 2.1.6 are not repeated here.

2.2.3.1 Percentage of Road Centre

The percentage of eye fixation on the road centre is probably the most commonly used parameter in the research for drivers' workload, in-vehicle system assessment, road condition evaluation, and driving performance. Many studies have shown that drivers look to the road centre region (at a few degrees down from the horizon gaze angle) about 80–90% of the time, and only rarely look at the lane markings near the vehicle during the normal driving (Harbluk, Noy & Eizenman, 2002; Recarte & Nunes, 2000; Rockwell, 1972; Wann & Swapp, 2001; Wierwille, 1993a, 1993b).

Victor et al. (Victor, Harbluk & Engstrom, 2005) have found that the percent of road centre viewing was decreased according to the increasing difficulty of the in-vehicle visual task, while it was increased during the auditory task. The sensitivity of gaze concentration towards the road centre to the secondary tasks was also found (Jamson and Merata, 2005; Engstrom, Johansson, and Ostlund, 2005).

This parameter is generally taken as "the higher the better" based on the assumption of "looking away from the road ahead at any time is harmful". However, this does not always hold true. For example, when the road condition itself is less demanding, drivers may have spare resource to look into peripheral areas more often. Another example is that compared to novice, an experienced driver looks less at the road ahead (Crundall, Underwood & Chapman, 1999a; Underwood, Crundall & Chapman, 2002), because they can use their peripheral vision more efficiently and be more selective in the mirror checking. Therefore, this parameter needs to be considered in relation to driving situation and task condition.

2.2.3.2 Total Number of Times of Fixation on Display

The total times of fixation is the total number of fixations made to the in-vehicle display during a visual task, where each fixation is separated by at least one glance at a different target. This measurement shows the number of times that drivers need to look at the display in order to perform a specific visual task. This measurement is widely used and intensively debated in safety research and in-vehicle system design. Initially, accumulated fixations of more than 9 times on an in-vehicle display, were suggested to be *unsafe* (Hankey et al., 2000). Later Blanco et al. (Blanco et al., 2006) investigated many of the visual tasks, and found actually almost half of the participants have to make more than 9 fixations at most displays to perform each task. Therefore, some other research suggested that the total number of fixations is not important, and even values greater than 10 are acceptable and routinely used (Chiang et al., 2004), and they suggested that it is more important to guarantee the task to be "chunkable" so that it can be finished in separate incremental steps, as long as each glance duration is typically less than 2 seconds.

2.2.3.3 Accumulated Fixation Duration

Accumulated or total fixation duration is the total amount of time in which fixations are made on the in-vehicle display. It is a parameter which suggests in total, how long a certain amount of information takes to be extracted by the driver. Accumulated fixation duration is an efficient measure which reflects how complex and demanding a given visual task is.

2.2.3.4 Total Task Duration

Total task duration is not an eye movement measurement, but it is listed here because of the tight correlation with visual searching behaviour. A shorter task time is desirable, as suggested by previous studies (e.g., Hashimoto and Atsumi, 2001). There were a number of generally accepted proposals. Green et al. (Farber, 2000; Green, 1999a, 1999b) proposed a "15-second rule" for in-vehicle navigation systems, which suggested that a Human-Machine Interface (HMI) task should not be any longer than 15 seconds. Later, Green suggested that to minimize the risk of distraction/overload, the criteria need to be dropped to 10 seconds (Green, 2008). Other researches argued a much longer duration, for example, Chiang (2004) suggested that acceptable total task times can be in the range of 30-40 seconds, or greater. Nevertheless, it is still generally believed that the longer task duration has a negative effect on workload and driving performance (Angell et al., 2006).

2.2.3.5 Fixation Duration

Fixation duration is the measure of interest when evaluating in-vehicle displays. A study by Chiang et al. (Chiang, Brooks & Weir, 2004) showed that, when involving in manual input, drivers' mean fixation duration on a display is one second when a single keystroke is entered, and about 1.5s when 2 keystrokes are entered in one *chunk*. Also, 94% of all display fixations in the cases of 1 or 2 keystrokes are less than 2s. It was therefore suggested that the fixation durations in the secondary task should be typically less than 2s (Chiang, Brooks & Weir, 2004). When a decision making process is involved, a maximum single fixation time during a task can be longer, about 2.5s (Blanco et al., 2006). At the same time, during the destination entries, the road was fixated for an average of 0.47s each time, with 95% of the roadway fixations being less than 1.2 s (Chiang, Brooks & Weir, 2004).

2.2.3.6 Percent Fixation Durations Exceeding 2 Seconds

Based on the research on the fixation duration, it was believed that durations longer than 2s have an impact on safety and therefore needs to be specially investigated. It was found that, when the in-vehicle display is more complex, drivers' percent fixation durations which are longer than 2 s increased (Victor, Harbluk & Engstrom, 2005). This measurement might be more significant than fixation duration solely. However, the *2s threshold* needs to be more carefully proven.

2.2.3.7 Standard Deviation of Fixation Duration

Standard deviation of glance duration is the standard deviation of the durations of all fixations made to a display during a visual task. Victor, Harbluk, and Engstrom (2005) investigated several eye-movement measures, including the mean fixation duration, the standard deviation of glance duration, the glance frequency, the total glance duration, and the percent on the road centre. In all of these measures, the standard deviation of fixation was found to be more sensitive, more robust, more reliable, and easier to calculate than the established glance-based measures. The standard deviation of fixation duration is increased with the increasing difficulty of visual task, which means looking at the display for more varied durations. It was found that a similar increase occurred during auditory tasks.

2.2.4 Summary of Eye Movement Review

As eye movements plan every action sequence, and organises activities of human behaviour, it could be a very sensitive parameter which reflects real-time or *instantaneous* behaviour. There is a close relationship between where we look and what we do. In driving, drivers must simultaneously keep the car in lane, avoid other traffic and be aware of road signs. In doing this, the eyes provide crucial information at all times. Some eye movement parameters can assess drivers' workload, for example, the blink duration and frequency in a mental task, and some eye movement measurements (e.g. the objects that drivers look at and the duration) reflect the distribution of attention and suggest the level of difficulty of information being processed. All these findings suggested the potential of eye movements to help understand drivers' coping strategies and behaviour change, which cannot be discovered by traditional performance measures alone. However, it needs be borne in mind that driving is a complex task with many elements which could impact on drivers' visual behaviour, and the eye movement itself is very sensitive and task-specific. Therefore there are two preconditions for effectively using eye movements in this research:

- Visual behaviour has to be considered relative to different task types and driving circumstances. Each eye movement parameter has to be well understood under different task condition;
- Future eye movement research is expected to reveal more of the cognitive processes that are hidden behind eye movement patterns for driving.

2.3 The Effects of In-Vehicle Tasks on Driving Performance

Performing in-vehicle secondary tasks whilst driving is related to distraction or overload, and eventually may cause driver errors (Carsten & Nilsson, 2001). Mobile phone conversation is a classic example of drivers engaging in in-vehicle task, which has stimulated intensive research. A study carried out by the Transport Research Laboratory TRL, UK, using a driving simulator (Burns et al., 2002) showed that driving performance when using a mobile phone was significantly worse than that under the influence of alcohol of the legal limit in the UK (Blood Alcohol Content, BAC: 80mg/100ml), as suggested by the reaction time of drivers' to a signal. Today increasing number of countries has banned the use of hand-held mobile phones whilst driving, but using hands-free phone is still legal in most countries. However, later studies have shown that hands-free mobile phone speech while driving also degrades coordination and control (Matthews et al., 2003; Nunes & Recarte, 2002; Tornros & Bolling, 2005). IVISs related in-vehicle tasks have similar effects. Rakauskas (Rakauskas et al., 2008) compared drivers performance while dealing with in-vehicle tasks (e.g. adjusting radio) and that of when they consumed alcohol (target BAC 0.08g/dl). The results suggested that performing in-vehicle task caused higher performance impairments than consuming the amount of alcohol above legal limit, as suggested by both longitudinal and lateral control of vehicle. For example, the standard deviation of lane position when performing in-vehicle tasks increased more significantly from about 25.25cm in baseline to 20.5cm, comparing to 17cm under the alcohol condition.

The design of HMI (Human Machine Interface) has a major effect on workload and driving performance. This includes: *design and location of buttons, controls and screens (size, brightness); menus; means of dialogue between the user and the system; the channel for information exchange (auditory, visual or haptic)* (Carsten & Nilsson, 2001). In the European Statement of Principles (European Commission DGXIII, 1998), it is stated that:

- *Visually displayed information should be such that the driver can assimilate it with a few glances which are brief enough not to adversely affect driving.*

- *Information relevant to the driving task should be timely and accurate.*
- *The system should not present information which may result in potentially hazardous behaviour by the driver or other road users.*

There are some general-accepted criteria for the design of IVIS in terms of driving performance "break down". Green and his colleges (Farber, 2000; Green, 1999a, 1999b) proposed "15-second rule" for in-vehicle navigation system, which suggested that HMI task should not be any longer than 15 seconds, which led to debate around the reliability of the task time and other criteria. For example, Green himself later suggested that to minimize the risk of distraction/overload, the criteria of task time need to be drop to 10 seconds (Green, 2008). Hankey et.al (Hankey et al., 2000) proposed that the safety issue is more about the number of separate glances towards a display, rather than the entire task time. He suggested limiting it to be less than nine glances. This is supported by the finding that drivers have the ability to divide a large task into small sub-tasks of between 1 to 2 seconds glace duration (Wierwille et al., 1988). Carsten (Carsten & Brookhuis, 2005) also questioned the correlation between task duration and visual workload since various levels of visual demand caused by different tasks. A significant amount of research was conducted to investigate the effect of in-vehicle tasks on driving performance, for example, when using in-vehicle navigation device (Ma & Kaber, 2007; Walker et al., 1991), traveller information system (Liu, 2001), in-vehicle E-mail system (Harbluk & Lalande, 2005; Lee et al., 2001), in-vehicle entertainment system (Horberry et al., 2006), reading in car electronic map (Tsimhoni & Green, 2001), and reading in-vehicle text (Hoffman et al., 2005).

When evaluating the effects of workload on driving performance, it needs to bear in mind that visual and mental load have distinctively different effects. In HASTE (Human Machine Interface And the Safety of Traffic in Europe), an EU FP5 project, a set of surrogate IVIS tasks was used to evaluate the potential distraction of an in-vehicle information system on driving performance both in simulator and field test (Carsten & Brookhuis, 2005; Engström, Johansson & Östlund, 2005; Östlund et al., 2004). Higher visual workload was related with more steering wheel adjustments and higher standard deviation of lateral position, i.e. more swerving in the road; and accompanied by a speed reduction and an increase in headway. This phenomenon of a speed reduction when engaged with an IVIS has been noted in numerous studies. It is generally believed that more complex visual tasks cause more deteriorated driving behaviour, because long text task more processing resources (Labiale, 1996), and the number of lines in the text display increase mean glace duration and number of glances greater than two seconds (Hoffman et al., 2005), which contribute to worse lane control. The effect of visual tasks is mainly due to looking away. Tsimhoni and Green (Tsimhoni and Green, 2001) compared driving performance between condition of reading an electronic map and “not looking at the road some time” – using occlusion technique. The results showed that that adding a secondary task significantly degraded driving performance, and the effect is similar to the occlusion condition. Drivers wandered more in their lane and departed from their lane more frequently, and the increase in the standard deviation of lane position was approximately 80%. Minin et al. (Minin et al., 2011) examined the impact of interacting with a visual searching task with two levels of difficulty in a simulated lane-change task. The results suggested that lane deviation indicated by three measurements was significantly affected not only not only between single and dual task conditions, but also between interacting with an easy and a difficult secondary task.

On the other hand, the cognitive task caused reduced lateral deviation, which was explained as a tendency for drivers to compensate for the task load by shifting away from the road edge, accompanied by an increase in glance focus on the roadway straight ahead and less on the periphery, so called “cognitive narrow”. While cognitive demands of secondary tasks increase, drivers tend to concentrate more on secondary task rather than peripheral or primary task. The decreased reversal rate while performing auditory tasks were also found in Brookhuis’s study (Brookhuis, De Vries & De Waard, 1991). Similar with that of visual tasks, complexity of auditory tasks also causes different performance. While drivers involved in a complex phone call, their reaction time to the collision warning is much longer than that in a simple conversation, especially when the warning is also given in auditory tunnel rather than by tactile (Mohebbi et al., 2009).The effects of the cognitive task on speed were not consistent across the studies (Carsten & Brookhuis, 2005; Engström et al., 2005; Östlund et al., 2004). It was therefore generally believed that the auditory task has less negative impact on drivers’ behaviour than the visual distractions, which further confirmed the MRT theory (Wickens, 1987a). In a review of speech-base in-vehicle systems, Baron (Baron & Green, 2006) summarized the researches in this area that: *Speech interfaces led to less workload than manual interfaces and reduced eyes-off-the-road times. Task completion time was generally less with speech interfaces, but not always (as in the case of manual phone dialling)*. However, this may not always hold true, nor does this suggest that auditory distraction would be tolerable. It was found in a field test that the inappropriate driving behaviour of interacting with pedestrian when conducting auditory tasks was NOT less deteriorated comparing to visual one, as suggested by delayed response or lack of response to the situation, auditory tasks caused even higher negative effects (Östlund et al., 2004). In another research investigating the effect of speech-based E-mail system carried out in a simulator, a 30% increase in reaction time to the leading vehicle brake was found (Lee et al., 2001).

The location of an in-vehicle display is another impact factor which contributes to the performance changes. Traditionally, the display is installed near the control panel, known as head-down display (HDD). In recent years, some vehicle manufacturers introduced in-vehicle head-up display (HUD) to deal with the negative impact because of the long “eye off road” (e.g. Nissan Toobe® and Ford e-Car®). Liu (Liu & Wen, 2004) investigated attention demands and driving performance while interacting with HUP in a simulator. The results showed that subjects reacted faster to the changes of speed limit signs, with lower steering wheel variation and smaller lateral acceleration than control condition. However, a pervious research comparing HUD and HDD in a driving simulator did not find any difference either in primary driving or in secondary task performance (Hooey and Gore, 1998).

Touch-screen not only requires looking away from the road ahead, but also involves manually operation; therefore raises even higher workload while driving comparing to traditional in-vehicle systems. A previous study by Paelke (Paelke, 1993) illustrated the detrimental effects of destination selection via several touch-screen input devices on a simulated driving task for both elderly and younger users. Tsimhoni et al. (Tsimhoni et al., 2004) compared the impact of two speech recognition interfaces and a touch-screen interface on driving performance in terms of total task time, standard deviation of lateral position during entering addresses into a navigation system during driving in a simulator study. An increase of standard deviation of lane position by 60% while interacting with touch-screen interface was found. Another research intended to investigate the user experience on four types of touch-screen feedback (visual only, auditory and visual, haptic and visual, and a combined feedback of auditory, haptic and visual) was carried out by Serafin and colleagues (Serafin et al., 2007) in both USA and Germany. The combined feedback was found to be the most preferred. Similar result was found by Pitts (Pitts et al., 2009). Fuller and colleagues (Fuller et al., 2008) investigated the effect of the location of touch-screen placement in a driving simulator. In this study, car-following was assigned as primary task while interacting with a menu-based interface on a touch screen located in different locations in the vehicle as a secondary task. The results showed that greater visual distance of the touch-screen from road ahead, or further away from subjects resulted in longer task duration. However, driving performance, as measured by RMS error (between subject vehicle and leading vehicle) and delay in speed changes, was not affected by the location of the touch-screen.

The format of how information is presented also has effect on driving performance. Blanco (Blanco et al., 2006) found that tasks presented in paragraph format were generally harder to be safely completed comparing to formats of table and graph. Between table and graphic display, it were found that the table format was more suitable for the tasks which involved computation elements, while the graphic format was better when there were only search and decision making elements required. When information is shown in text, as found in another study (Hoffman et al., 2005), the line by line display is more demanding than page by page, particularly when drivers have to manually scroll down to check the next line.

However, it should be noted that the effect of IVISs can not always be detected by driving performance alone, especially for the auditory ones. It was suggested that in the assessment of IVISs, driving performance may not be the only determinate element, because IVIS distraction may cause higher mental or visual workload without necessary impair performance due to the drivers' extra effort (de Waard, 1996). Sometimes, although the temporal performance still kept in an appropriate level, the stress from longer or frequent exposed to high workload may eventually cause the performance decrease (de Waard and Brookhuis, 1997). A multidivisional matrix which includes subjective mental workload measure, physiological measurements and driving performance is a promising solution.

2.4 On-Road Test versus Driving Simulator

Driving performance can be evaluated either in an on-road test or a driving simulator. The on-road evaluation is considered as the standard test approach and most ecologically valid method for assessing drivers' fitness of driving (Fox et al., 1998; Griffen et al., 2011; Reger et al., 2004). It was stated that when assessing the effect of overload or distraction from a system, the only way is by empirical experience in real world and see how drivers will respond to it (Carsten and Nilsson, 2001). On-road experiments can revealed the "ground truth" of driving behaviour, and the effects observed under these circumstances are more reliable comparing to the results from simulator (Carsten & Brookhuis, 2005; Santos et al., 2005). It was believed that even short-duration on-road evaluations can reflect real-world driving situations, and are therefore used often to examine the effects of performing in-vehicle tasks (Cooper & Zheng, 2002; Green, 1999a; Harbluk, Noy & Eizenman, 2002). By collecting data of vehicle trajectory and manoeuvring using a instrumented vehicle, driving performance can be quantitatively assessed in an on-road test. This method is free of human bias and can observe drivers' behaviour as close as that in reality (Östlund et al., 2004). Most instrumented vehicles provide continuous data stream of vehicle speed, brake pedal pressure, throttle position, engine speed and steering behaviour. Lane keeping, relative distance to adjacent vehicles, merging and following behaviour of a driver can be gathered through radar and video systems (Guensler & Willems, 2002; McDonald & Brackston, 1997). Hazard behaviour like abrupt brake, large and sudden steering wheel adjustment, or swerving on the road can therefore be detected by setting specific criteria on related vehicle measurements. The data captured from instrumented vehicles were broad used in transportation research, for example, to identify factors affecting risk of accident, investigate travellers' behaviour, to assess the design of vehicle and new systems, to evaluate the effect of advanced travel information, to detect performance deterioration caused by fatigue, etc.. In an on-road test, drivers are exposed to the real traffic and risk, especially when we introduce extra workload to them. To ensure safety, a co-driver or a safety observer is often presented in a test vehicle to accompany the driver, and also, the secondary task complexity should be designed in a lower level of complexity (Östlund et al., 2004).

Blanco and colleagues conducted a series of three on-road experiments to investigate the secondary tasks from an IVIS on driving performance both in passenger car and in truck (Blanco et al., 2006). 36, 12 and 36 participants drove along a 4-lane divided section of U.S. following a leading vehicle. While driving, they were presented with IVIS tasks with various information densities, decision-making elements, presentation formats, and presentation modalities (visual or auditory). The results showed that, for both presentation modalities, the presence of a secondary task had a substantial negative impact on driving performance of both automobile drivers and truck drivers. The authors stated that it is critically important to consider not only the design of the display, but also whether and to what degree the tasks will be performed. An on-road experiment carried out in real traffic is capable of capturing drivers' behaviour in real world. As suggested by Horrey and Lesch (Horrey & Lesch, 2009), as the true "risk" in real traffic could not be replicated on a closed test track, drivers' strategy of performing secondary tasks may not reflect what they would do in reality, for example, they may have been more willing to engage in in-vehicle tasks on test track, or in a simulator. Schrauf and colleagues examined the impact of auditory secondary tasks on driving performance in an on-road experiment in German (Schrauf et al., 2011). It was found that the brake reaction time to a leading vehicle increased significantly when drivers engaging in the auditory tasks, indicating reduced driving performance. The electroencephalographic (EEG: alpha spindles, alpha band power) and cardiac activity (ECG: heart rate variability) were also collected in this study, which showed similar effects, i.e. Alpha spindle rate, alpha band power as well as heart rate variability (HRV) increased with time-on-task and were significantly different during the secondary task, indicating inhibited visual information processing and reduced concentration ability. The authors suggested that a coherent effect from the EEG, ECG and brake reaction time enables the direct quantification of driver distraction in experiments during real road driving.

Laboratory setup can indicate the impact of IVIS on driving performance to some extent, but it is unable to provide a sizeable influence for the different levels of task difficulties (Santos et al., 2005). While a simulator can provide some advantages, for example safety (Kaptein et al., 1996) and the ability to offer better experimental control, the effects of increased workload on driving performance can be exaggerated using simulators (Reed & Green, 1999; Santos et al., 2005) compared to those achieved from actual driving. For example, the variation in lane position (Reed & Green, 1999) can be twice as large in the simulator as in an instrumented vehicle, which may reflect drivers' tendency to be less cautious about making errors in the simulator, as the consequences for doing so are far less than in real driving. Another draw back of simulation research is the Simulator Adaption Syndrome (SAS) caused by the mismatch between visual cues of movement and inertial cues. The symptoms of SAS include dizziness, nervousness, light-headedness, body temperature increase, and nausea, and women and older drivers were more likely to suffer from the SAS. Furthermore, the artificial environment of a simulator or laboratory could also lead to *different* strategies from drivers when dual-tasking, as drivers tend to *prioritise* secondary tasks over primary driving, given there are no consequences arising from errors (Goodman et al., 1997). Carsten and Brookhuis (Carsten & Brookhuis, 2005) stated that most simulators are not capable of capture drivers' behaviour when the primary driving becomes really cognitively demanding, i.e. where more understanding and interpreting of the on-road situation are required, because only simple driving-related cognitive tasks are tend to be used (e.g. understanding a road sign). In the HASTE project, some field tests were conducted, which include a straight four-lane motorway section outside of Linkoping, Sweden. The results showed that the field studies tended to pick up somewhat different effects of the systems than the simulator studies (Östlund et al., 2004). For example, the study in the simulator (Jamson & Merat, 2005) showed that drivers compensated for both visual and auditory tasks by reducing their speed, and this was more prominent during interaction with the visual task. However, this effect of auditory tasks was not found in field test.

"Naturalistic Driving" Study is another method to access drivers' performance in real traffic. In this method, drivers are not given special instructions, no experimenter is present, and the data collection is realised by pre-installed on-board instrumentation, which is unobtrusive to the driving (Cheng et al., 2011; Dingus et al., 2006). The current technology also allows drivers to use their own vehicles, with only some software and hardware assembled before the data collection process. One of the advantages of conducting naturalistic driving study is to get accessible to extreme cases of driving behaviour and performance, including severe drowsiness, impairment, judgment error, risk taking, willingness to engage in secondary tasks, aggressive driving, and traffic violations over a long period of time. Although this method is suitable to analysis accident related crashes and near-crashes events, it is not suitable for the current study, because the difficulty of standardise all types and levels of workload encountered in real life. While in a pre-designed experiment, we can manipulate the task demand by controlling the road-context, time of tasks being presented, types and difficult levels of secondary tasks to create a scenario in which drivers' errors and behaviour changes can be observed and evaluated

(Rizzo et al., 2001). Also, it was stated that the analysis on real-life accidents is also not credible for assessing driving performance, because accidents are sporadic events that are uncontrollable and difficult to be objectively evaluated (Rizzo, 2004). Therefore the "Naturalistic Driving" method which utilise instrumented vehicle was not applied for this research.

Driving performance can also be assessed by other subjective method, for example, by examiners and assistants (Mayhew et al., 2011), or a accompanying observer (Anttila & Luoma, 2005). However, these methods are subjective, and the opinions between different examiners are not likely to be consistent.

For reasons mentioned above an on-road experiment was considered as the most valid and reliable method for conducting research into the impact of IVIS secondary tasks on driving performance. When design an on-road experiment, the challenge is that fluctuation in traffic and road condition could reduce the strength of the correlation between driving performance and the neuropsychological measurements (Rizzo, 2005). The design of such experiments can be more complex, e.g. to account for extra "noise", and the data processing and analysis can be more difficult and laborious. Therefore, systematic consideration of driving scenarios in which on-road tests to be conducted is of importance, for example, a relatively consistent traffic flow needs to be as consistent as possible. Previous practice includes conducting experiments on urban road (Anttila & Luoma, 2005; Tornros & Bolling, 2006), free driving on rural road (Jamson & Merat, 2005; Tornros & Bolling, 2006), in intersections(Caird et al., 2008; Rizzo et al., 2001), curvy road section (Hong et al., 2006; Marple-Horvat et al., 2005; Sodnik et al., 2008)) , car-following on a rural or dual-carriage road section (Brookhuis et al., 1994; Schrauf et al., 2011), performing lane changing (Mattes, 2003; Rognin et al., 2007), and motorway driving (Engström, Johansson & Östlund, 2005; Young et al., 2009). Lessons can be learnt form the experience from HASTE project, as one of the intentions was to single out the most powerful scenarios for evaluating the effect of IVISs on driving performance (Östlund et al., 2004). Motorway driving in the various simulators and the laboratory tests were found generally less diagnostic, than driving on other road types. While the rural road was suggested to be the most diagnostic, i.e. the effect sizes from the rural road were generally larger. The urban road did not pick up any additional information that was not provided by the rural road. This means that, for simulator and laboratory assessments, the rural road can be used as the sole road category in the later work of HASTE assessing real IVIS systems as well as in the final HASTE test procedure. In the field studies with the cognitive task, the motorway produced the only indicator with a consistent effect. Another field test was completed in a build-up area in Linkoping, Sweden in order to study the possibility of using Peripheral Detection Task (PDT) as a standard method for safety testing and evaluation of IVIS (Harms & Patter, 2003). No significant performance decrease was found in that study, as measured by speed and brake force, and only a negative impact of using of navigation system on PDT performance was found. This again suggested that urban route may not suitable for this study. In an on-road study carried out in the US, Blanco et al. (Blanco et al., 2006) tested the effect of IVIS-related cognitive process on driving performance by a study in both passenger car and lorry on 4-lane divided section of major highway (Route 460 in US). In this study, drivers were asked to perform searching or decide to compute a best route to destination according to the information provided on in-vehicle display, which are the decision-making tasks. A substantial negative impact was found when there were multiple decision-making elements involved compare to the one-decision-making element in conventional tasks. There were several other on road test on in-vehicle information system designs carried out in USA (Hankey et al., 2000b, Blanco et al., 2006), Finland , and Netherland. Design guidelines and in-vehicle display design safety criteria were proposed accordingly. However, since the rather different driving situation, the usability of these criteria in UK is not clear.

In summary, to help with which experimental method should be employed in this study, the lessons learnt from previous research on the effect of distractions, in-vehicle tasks, and other factors on driving performance suggest:

- After consider a number of possible methods, an empirical observation in real traffic conditions using an instrumented vehicle should be used to understand drivers' actual behaviour changes and coping strategies for the objectives of this research.
- The effects of auditory and visual workload needs to be investigated separately, due to the opposing effects of the two, and therefore any workload containing *both* auditory and visual elements will cause complicated effects and are subject to too many uncontrollable factors.
- The use of driving behaviour and eye movement parameters is subject to context, e.g. the type of secondary task being performed, and under what driving circumstance. Therefore, driving scenarios in which on-road tests are carried out need to be carefully investigated and selected to provide a relatively stable traffic flow, as well as be "diagnostic" enough to differentiate the extra workload exerted and their associated performance levels. This includes considering the likely traffic flow on e.g. a truck road or motor way, the geometric layout of the road (e.g. whether straight, curvy or bumpy) and the need for consistency. Other factors, for example time-of-day and traffic variations, also must be taken into account in the experimental design. As there is insufficient guidance found in the literature for the design of such high workload research in real traffic on UK roads, a series of pilots need to be run using different road

conditions to select the best scenarios for research, although experience from other countries can help in decision making.

2.5 Conclusions from Literature Review

From the literature review, various *descriptive* workload and performance models have been proposed which suggest that, whilst driving, a driver's workload is increased by additional task demands: when the workload is relatively low, drivers can deal with the increase by using different coping strategies. However, when task demands (and therefore workload) increase above a certain level, the driver's performance will decrease. Therefore, with these models, the *quantitative* changes between an increase in workload and a reduction in performance are still *unclear*, due to the lack of a recognised, sensitive, instantaneous, and non-invasive approach to assess workload. Also, the *extent* to which drivers' coping strategies can impact on the relationship between workload and driving performance can rarely be observed by traditional measurements. Given that these coping strategies differ from individual to individual, and some strategies are more effective than others (or require fewer resources to engage), they consequently have different impacts on the ability to maintain driving performance. Therefore, *existing* models which describe the relationship between workload and driving performance are *too simplistic*, and are not adequate for transport applications.

A review of the literature on visual behaviour provides good evidence for the potential use of eye movement measurements as a more effective tool for assessing workload, and to evaluate its impact on performance. By combining attributes of both psychological *status* and *behaviour*, eye movements provide a set of sensitive, diagnostic, and instantaneous measurements for relating cognitive status and behaviour, which can potentially be used to provide an improved understanding of the relationship between workload and driving performance, and to assist in exploring drivers' associated coping strategies.

The reviews on the previous findings on the effects of in-vehicle tasks on driving performance and the comparison between different research approaches provide the expected results and the "best practice" for this study. It was therefore decided that an empirical observation in real traffic conditions will be used for the objectives of this research; the effects of auditory and visual workload will be investigated separately, and a series of pilots using different road conditions are need to select the best scenarios for research, although experience from other countries can help in decision making.

Chapter 3 The On-Road Experiment

3.1 Introduction

Following the literature review, it was decided that an empirical experiment would be the most suitable for this study, i.e. to assess drivers' performance from increasing workloads due to the addition of *secondary* in-vehicle tasks. As discussed in Sections 1.2 and 1.5, the objective of this research was to investigate the impact of these tasks on drivers, and how they cope with the increased workload whilst driving and performing in-vehicle tasks simultaneously, i.e. dual-taking. In order to achieve this, real traffic conditions were required to observe drivers' on-road behaviour changes and coping strategies. This requires the collection of their eye movements or visual behaviour with their actual vehicle control and manoeuvring data, which cannot be obtained through other methods, including simulator, using a test-track, subjective surveys or accident databases. Therefore an On-Road Experiment was conducted using an Instrumented Vehicle, where "live" on-road data would provide a real and improved understanding of the relationship between secondary task-induced workload and driving performance, as well as individual driver behavioural changes, for example in their eye movement and attention shifting. (See Section 2.4 also for a discussion of why other methods were discounted.)

In the experiment, drivers of different ages, genders and experiences were recruited to participate in the on-road experiments. The Transportation Research Group (TRG)'s Instrumented Vehicle (IV) was used to run field tests in two of the most-common driving scenarios: Car-Following and Free-Driving. During the tests, drivers were required by the experiment coordinator ("the experimenter") to perform a set of pre-designed secondary auditory and visual tasks whilst driving. The auditory and visual tasks were separated to investigate their distinct effects on individuals caused by these two different types of workload (Östlund et al., 2004). These secondary tasks were presented to drivers using software or *Operation Simulation System* developed specifically for this study. As well as the *driving* behaviour data recorded by the IV and this System, the FaceLAB™ *Eye Monitoring System* was also used to record drivers' *visual* behaviour. After each driving session, the drivers were interviewed about their experience of the secondary tasks and self-perception of any behaviour changes. Subjects were told that the aim of the study was to investigate impacts of in-vehicle secondary tasks. The objective of study on their coping strategy was not revealed to them. This is to prevent subjects from knowing the information that might lead to conscious or subconscious bias, which may further cause invalidate results.

The overview of the experimental design is given in Section 3.2 below. This is followed by the detailed design of the driving scenarios and secondary tasks in Sections 3.3 and 3.4 respectively. Details of how the participants were selected for the experiment, the apparatus used, and the test routes selected for the driving scenarios are given in Sections 3.5, 3.6 and 3.7. The questionnaires and surveys used for the experiment/interviews are described in Section 3.8, followed by the experimental procedure in Section 3.9. These Sections collectively form the methodology and design of the On-Road Experiment used for this research.

3.2 Experimental Design – Overview

The experiment is a *mixed* design of *within-subject* and *between-subject* factors. From the literature review, a set of factors highly relevant to drivers' behaviour when distracted was identified, and the effects of these investigated using the experiment. These comprise:

- within-subjects factors:
  - *driving scenarios*;
  - *secondary task types*;
  - *task difficulty* (or complexity);
- between-subjects factors:
  - *gender*;
  - *age* and *driving experience;*
  - *experience* and *attitude towards using IVIS*.

The different within-subject and between-subject factors are discussed separately in Sections 3.2.1 and 3.2.2.

3.2.1 Design of Within-Subjects Factors

The within-subject factors comprise the driving scenarios, and the secondary task types and associated levels of task difficulty, which are considered below.

3.2.1.1 Driving Scenarios

The impact of IVIS distractions on driving performance cannot be assessed using one scenario alone, as the measureable parameters (such as speed and headway), can also be impacted by the driving conditions and circumstances. The experiment therefore needs to be performed using more than one scenario, to compare the *consistency* of performance effects across different scenarios, and to observe whether these effects are indeed caused by dual-tasking (i.e. they are due to an increased workload from the IVIS task). To address the problem of noise from real-road experiments, including uncontrollable traffic, road and weather conditions, the scenarios for the experiment were carefully selected to ensure these effects were minimised, and the experimental conditions were as consistent as possible. After an investigative series of pilots over several months, the following two scenarios were selected as the most suitable for this experiment:

- a *Car-Following* scenario in the inner lane of a dual-carriageway trunk road between Winchester and Newbury, and
- a *Free-Driving* scenario on a very quiet section of a straight A-road between Winchester and Basingstoke.

In these two scenarios, it was seen from the pilot studies that drivers' behaviour was *least* affected by the traffic conditions and other external factors. Further details of the driving scenarios, as well as the test routes used, are given in Section 3.3 and 3.7.

3.2.1.2 Secondary Tasks - Type and Complexity

Previous research suggests that *mental* and *visual* tasks have very different effects on driving performance (Anttila & Luoma, 2005; Engström, Johansson & Östlund, 2005; Jamson & Merat, 2005). A series of artificial secondary tasks was therefore used to *simulate* the effects of using IVIS to induce this extra workload on drivers. These tasks are easier to control in terms of where, when and how they are presented to drivers, with the ability to separate *distinctively* the mental and visual demands, and the level of difficulty exerted to drivers. The drivers were therefore asked to perform two different *types* of secondary tasks, auditory and visual, to differentiate between the effects of mental and visual distractions, and the workload they caused. The tasks were specifically designed to investigate the effects of increasing workload on drivers' performance and visual behaviour by inducing varying levels of mental and visual demands on drivers through different levels of difficulty (*demand* levels 1, 2 and 3). For each level of difficulty, drivers were asked to perform the visual or auditory task three times, in both driving Scenarios. Drivers were also asked to drive normally (without performing any secondary tasks) for each Scenario, to act as a baseline for comparison. As had been suggested by HASTE (Östlund et al., 2004), the risk of accidents is primarily related to drivers' attention, visual behaviour, and performance. For safety reasons, the secondary tasks used in this experiment were designed so as to induce the extra workload required, but were *not* so high as to cause accidents. Further details of the secondary tasks used in the experiment are given in Section 3.4.

3.2.2 Design of Between-Subject Factors

The between-subject factors cover a range of different *driver characteristics* including gender, driving experience and age, and experience and attitude towards using IVIS. In carrying out the experiment, a range of participants were selected to cover the different groups of these between-subject factors, as ideally, there should be an equally distributed number of subjects across all the driver characteristics groups considered here, and also, within each characteristics group, the other impact factors should be as balanced as possible. Details of the people who were selected (or participated) in the Experiment are given in Section 3.5. From the Literature (Chapter 2), it was expected there would be some difference in the behaviour between men and women, and between those with difference levels of driving experience.

3.2.3 Other Design Considerations

In each driving Scenario, the *primary* driving performance and eye movement data for each driver was collected before, during and after the secondary tasks were performed. This before, during and after information was combined into a single dataset to form an *unbiased* baseline for each driver, to balance out the effects of familiarisation with the vehicle. Drivers' performance and eye movement data in conducting the secondary tasks were then compared against this baseline dataset to establish any behaviour change. To ensure that drivers were not more familiar with the IV in one Scenario versus the other, half the drivers performed the Car-Following Scenario first, while the other half performed Free-Driving first.

To reduce the learning effects from repetition of the secondary tasks, the order of performing the different task types and difficulty levels was counterbalanced across times of task repetition. In addition, a comparison of the first, second and third time when drivers performed a secondary task was conducted after the experiment to account for the effects of repetition, should they exist.

A comparison of the effects of different leading and tailing vehicle types on driving performance was also conducted after the experiment, to avoid their potential interaction with the impact of performing secondary tasks and the different driver characteristics. Where an effect exists, as was the case with Car-Following a van for example, the individual task cases were excluded from the main analysis.

3.3 Design of Driving Scenarios

Several criteria were used for choosing the appropriate driving scenarios for the experiment, namely that they had to:

- be representative of typical driving situations;

- be conducted in relatively stable traffic conditions (e.g. not during peak hours), to minimise the effects of heavy and/or fluctuating traffic flows (and the extra workload this causes drivers);
- be easy to control, and as reproducible as possible,
- be conducted in relatively straight road sections, to reduce the need (and extra workload required) for steering wheel adjustments due to road geometry; and
- provide similar levels of difficulty or complexity for drivers, across the different scenarios.

A serial of potential scenarios in which the on-road experiment would be conducted was proposed according to the previous practice from literature review (see Section 2.4). To assess the proposed scenarios, an on-road pilot (or series of tests) was conducted using the IV driven by two separate experienced technicians from TRG, with an experimenter sitting in the car as observer. The simulated IVIS secondary tasks were presented to each TRG driver to evaluate the feasibility and suitability of conducting the experiment in the given driving context. The process included making observations of the suitability of the road/road conditions through discussions between the experimenter and the technicians, and taking into account (and reviewing) the video images and auditor recordings taken along the route, e.g. for different leading and tailing vehicle types. In addition to the data recorded by the IV, the secondary task performance, driving behaviour, and vocally reported experience of their workload (both quantitative and qualitative) were also collected, and the information analysed after the tests to help determine which scenarios were the most appropriate for the experiment.

Initially, several driving scenarios were tested, including:

- an urban single and/or dual-carriage way;
- different road intersections (e.g. t- and y-sections);
- a rural, hilly road;
- a rural, winding road;
- car following on a rural road;
- lane changing, and car following on a motorway; and
- (unrestricted) free driving.

The scenarios using an urban route, intersections, hilly or winding roads, or car following on a rural road were ruled out after the pilot, because the addition of extra secondary tasks were said to cause too high a workload for drivers, and could therefore be dangerous. The potential and/or effects of interacting with external factors, such as other vehicles, pedestrians, cyclists, etc. in these scenarios also make the results hard to compare and interpret.

The tests on motorways suggested these scenarios were also too complex and demanding on drivers, while difficult to reproduce consistency. Hence, after three months of on-road testing, two scenarios were chosen as the most appropriate for this study, i.e. Car-Following in the inner lane of a dual-carriageway trunk road, and Free-Driving on a quiet road section of a straight primary ("A") road. Both scenarios were considered to be typical and realistic driving situations. In the case of the Car-Following Scenario, the pilot was also conducted to help identify driving behaviour which was specifically due to the impact of different leading and tailing vehicles, which helped to filter out the "noise" caused by these effects in the experiment. This was also performed for tailing vehicles in the case of the Free-Driving Scenario, to help reduce the impact of these.

Primary driving performance in the Car-Following Scenario is typically measured through longitudinal and lateral stability, i.e. the ability of drivers to maintain a more stable headway and lane position (or lower lane deviation). In Free-Driving, speed stability and, to a lesser extent lane control, are the key indicators of driving performance. In terms of eye movement, drivers tend to focus on the road ahead while Car-Following (to maintain a consistent headway), while this is not so critical for Free-Driving, although information still needs to be gathered for the road ahead.

3.4 Design of Secondary Tasks

According to the Multiple Resource Theory (MRT) (Wickens, 1987a), the performance of a task involves three stages: *information extracting*, *processing*, and *responding*, and there are three correspondingly aspects of secondary tasks which can affect workload: *information modality* (e.g. visual or auditory), *working memory required* (e.g. verbal or spatial), and *response types* (e.g. manual or vocal). Since each type of *resource* (or the *ability* to process information) is limited, when "dual-tasking" (in this case, performing a secondary task while driving), the greater the overlap of the resources required for the two tasks, the higher the anticipated conflict, and therefore the higher *expected workload* caused by the tasks. In "normal" driving (i.e. performing the primary task), the *perception*, *processing* and *responding* resources are more emphasised on *vision* (to perceive the road scene), *manual* (to adjust the steering, accelerator and brakes) and *spatial working memory* (to judge relative vehicle positions). Secondary tasks which require *auditory perception*, *vocal responding*, and *verbal working memory* are therefore expected to have the lowest overlap with "normal" driving (in terms of the competition for resources), and are expected to induce a relative low workload; while the ones which involve *visual information perception*, *manual responding*, and *spatial working memory* are expected to have the highest overlap with the primary task of driving, and therefore are expected to cause higher drivers workloads. The secondary tasks were therefore separated into two types: *mental-vocal* and *visual-manual*.

Four criteria were used to select the simulated in-vehicle tasks to be performed for the experiment. Each task was required to:

- induce different types of workload (mental, auditory, visual, and manual);
- cause different levels of expected workload due to the task demand;
- represent commonly-used or potential In-Vehicle Information Systems (IVIS) tasks; and
- minimise the effects of individual differences.

Two *input-output* task types were therefore used to simulate the different types of workload imposed by common IVIS tasks. A third task type was used as a “reference task” in order to compare the performance in the input-output tasks against a “baseline” or the workload caused by a relatively simple and common IVIS task, i.e. turning on the radio. The two input-output tasks were:

- **Auditory** input and **Vocal** output (“**Auditory tasks**”). For this task type, drivers were given an auditory instruction, and then asked to respond (i.e. performing the task) vocally, for example to distinguish how many times a given sound appeared in an audio. No *extra* visual information input or manual output was required for this type of task.
- **Visual** input and **Manual** output (“**Visual tasks**”). For this task type, drivers were asked to search visually for certain target object(s) on a touch screen, and to pick them out manually, for example by “clicking on” the required item(s) with their finger.

The design of these two task types as well as the reference task, are discussed separately below.

3.4.1 The Auditory Tasks

From the literature, auditory tasks have commonly been used to research the effects of mental distraction on driving behaviour, for example while making conversation with a passenger or using a mobile phone whilst driving. However, it is difficult to maintain a consistent level of workload in the design of a conversation task, and most auditory tasks also contain other intelligent test materials (Burns et al., 2002), for example computation (Brown et al., 1969), spatial imagination (Recarte & Nunes, 2000), wording (Recarte et al., 2008) and/or negotiation (Parkes, 1991b). Research conducted by TRL also applied the Rosenbaum Verbal Cognitive Test Battery (Waugh et al., 2000), which contains Repeat Sentences and Verbal Puzzles Task, in which drivers were required to repeat a sentence, or to answer verbal puzzle questions as quickly as possible using their judgement and flexible thinking with their response time measured. However, all these tests to some extent evoke extra individual difference (e.g. computation effort), which needed to be minimised in the experimental design, to order to investigate the differences between subjects factors. More *concept-free* auditory tasks were therefore designed.

These auditory task have the same concept as the cognitive tasks designed by Jamson and Merat (Jamson & Merat, 2005), and were adapted from a visual version of the Continuous Memory Task (CMT) used by Veltman and Gaillard (Veltman & Gaillard, 1998), where a series of target letters (A, AB, ABC, ABCD) were presented sequentially with non-target letters (i.e. other alphabetic letters) on a computer screen. Participants were asked to press a button each time a target letter appeared, and the number of times each target letter appeared was counted separately, i.e. if the target letter or letters appeared twice, then the button had to be pressed twice. The level of task difficulty was increased by increasing the number of different target letters that must be counted, and the participants' workload was assessed either manually or vocally. Given these tasks were deemed suitable for assessing the effects of increasing cognitive (i.e. mental) workload on driving behaviour, an auditory version of these tasks was created by replacing the letters with different bursts of sounds (Östlund et al., 2004), so as not to cause any extra visual workload on drivers.

In these tasks, drivers were now played one or more *target* sounds (each lasting about 320ms in duration) at the beginning of the auditory task, followed by 15 similar *background* sounds appearing at 2 second intervals, which included the target sound(s). Drivers were asked to remember the target sound(s) and then recall how many times **each** target sound appeared in the following 15 sounds, and to report the total for each target afterwards. Three levels of task difficulty were designed, with level 1 having one target sound, level 2 having two targets and level 3, three target sounds. To help drivers optimise their attention, they were given an auditory cue of the task difficulty level beforehand, e.g. "Sound counting task 2-1", which stands for task number one, in difficulty level 2 (or two target sounds).

The total sound pool includes 32 different short sounds, including prompt sounds such as "ding", which is commonly used in IVISs, and other salient and common sounds found in daily life. 10 different target sounds were chosen from this pool by the experimenter for the auditory tasks, as the objective was to increase drivers' mental workload to *distinguish* (and count) different target sounds, and not to remember all the different sounds. Four pilot trials were carried out to improve the *design* of the auditory task; two were desktop–based, and two were conducted on-road:

- Desktop Pilot 1: Seven students (4 male, 3 female) were asked to rate on a scale of 1 to 5 how memorable and salient the 10 potential target sounds were. Those that were rated as high for memorability and low in variations were selected for the target sounds – comprising seven in total.

- Desktop Pilot 2: Based on the target sounds selected in the Pilot 1, two auditory tasks were designed for *each* difficulty levels (i.e. six in total) – one of the tasks in level 2 was selected to be a *reference* task, with a fixed workload value of 100. The same seven subjects were asked to rate their demand (or workload) of performing these auditory tasks, by comparing to the reference task, where higher value represents a highest workload. The subjects were interviewed afterwards, and their subjective workload used to further distinguish the task difficulty levels, i.e. level 1: mean = 86.79, SD = 29.93; level 2: mean = 106.07, SD = 32.24; level 3: mean = 161.79, SD = 59.99 – see Figure 3.1. (Because of the small sample size, no further statistical analysis was conducted.)

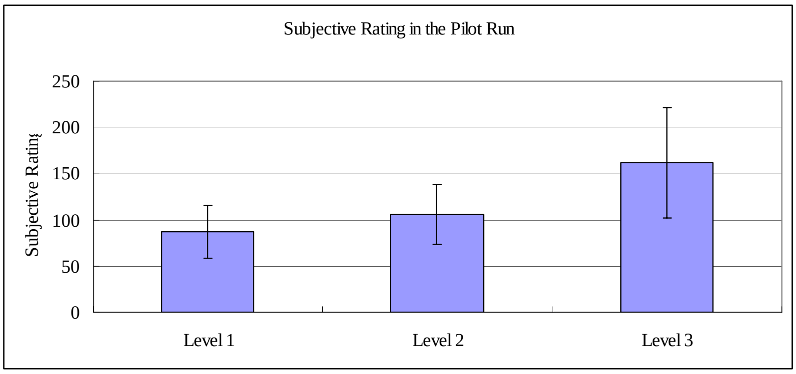


Figure 3.1 Subjective rating on the auditory task demand in the pilot run

The following issues were raised during the subject interviews:

1. it was hard to remember the three different target sounds used in difficulty level 3;
2. there was some similarity between the target and the background sounds used, which could cause confusion.

The target sounds were therefore cut further from seven different sounds to five of the most commonly found in daily life, for example the sound of a horn, in order to keep them significant and easy to remember. The background sounds which tended to cause some confusion with the targets were also replaced by other background sounds.

- On-road Pilot 1: An experienced TRG engineer (acting as a test subject) and the experimenter took part in an on-road pilot run of the auditory task, using the modified auditory tasks developed from the desktop pilots. All three task difficult levels were tested twice to verify the suitability and usability of the auditory tasks under real-road experiment conditions. The subject was asked to rate the workload when performing the auditory tasks against a pre-designed reference task, with a workload offset at 100, and taking into account the road condition and surrounding traffic. The subject was asked to report any issues encountered during the pilot run, which was conducted in parallel to exploring potential test routes. A few major issues arose during this pilot:

1. Due to the volume of noise on motorways, the sounds were hard to be heard, and therefore it was difficult to perform the auditory tasks;
2. The tasks were too difficult to perform whilst driving overall, so needed to be simplified;
3. The clicking sound of the laptop while operating the Simulation System software could be confused with the sounds required for the tasks, and therefore cause extra distraction;
4. Performing the auditory task in urban driving was also deemed too demanding and potentially dangerous.

After this pilot, issues surrounding the task design were raised, discussed, recorded, and the design adjusted and re-tested to resolve any problems. For example, a speaker was installed behind the IV's back seats to present the sounds more clearly, and the background sounds simplified to reduce their variety. The software was also re-designed to limit the number of clicks required to manipulate the tasks, and the auditory clips combined where possible, so that the distraction from having to click the pad on the desktop was minimised. Given it was hard and too demanding to hear the sounds in urban traffic and on motorways (and therefore potential dangerous), the use of these scenarios were excluded from the experiment.

- After correcting these issues, a final On-road Pilot run 2 was carried out to prevent any further problems which might arise during the experiment, but no further issues were found. Based on information collected in both Desktop and On-road Pilot, six auditory tasks for each task difficulty level were designed for the experiment.

3.4.2 The Visual Tasks

The visual tasks were designed to introduce extra workload to drivers visually on a touch screen, and required a manual response. Similar visual tasks have been used in previous driving behaviour research, for example, those involves map-reading (Burnett & Joyner, 1997), using navigation systems (Pohlmann & Traenkle, 1994; Uang & Hwang, 2003), understanding text (Hoffman et al., 2005), and colour reorganisation (Rosenbaum, 2000). Similar to the challenge of designing auditory tasks, a consistent level of increased workload between different subjects had to be maintained. Therefore a concept-free, more generalised design, which is representative of current IVIS systems, and which does not contain text reading, complex maps or other content subject to drivers' individual differences, was required for this study. The Arrow task, as used in the HASTE project was also considered, but was not adopted because this task contains additional spatial information (i.e. direction arrows), which would have subjected drivers to workloads that could have compromised safety.

Circle Tasks were therefore selected for use in this study, and the design of these tasks was based on visual search experiments frequently used in experimental psychology. According to Multiple Resource Theory (Wickens, 1984b), the combination of secondary visual tasks and driving are the most difficult to perform simultaneously, due to the conflict in visual and manual resources. Treisman's Feature Integration Theory (Treisman, 1988) suggests the speed at which a visual target is identified within a display is affected by its visual similarity to other objects in that display, in other words, that unique features of a target object allow it to 'pop out' of the display, resulting in faster decision times, while similar shaped and/or coloured objects makes them harder to distinguish.

In the visual tasks, drivers were asked to rapidly search visually *dense* scenes on a touch screen, to extract critical information from them – in this case, a target object, and to respond in a timely manner - in this case, by touching the object. The distractions, or non-target objects in the scene, can disturb this process in a bottom-up or top-down manner (Cole & Hughes, 1984; Engel, 1971; Hughes & Cole, 1986), and higher amounts of distractions require a greater number of fixations in order to locate the target (Boersema, Zwaga & Adams, 1988). Therefore, the difficulty in target identification increases as the non-target or background objects become more similar to the target in colour, shape and/or orientation. In addition, increasing the number of objects in a scene on the touch screen is shown to increase the reaction time to identify the target(s). Therefore three different levels of visual task difficulty were designed based on the number and the visual characteristics of the background objects.

The task itself is similar in concept to the one used in the ADAM (Advanced Driver Attention Metrics) project (Mattes, 2003), which has been used and validated by on-road experiments, for example Rognin (Rognin et al., 2007), Wynn and Richardson (Rognin et al., 2007; Wynn & Richardson, 2008) and Ranney et al. (Ranney et al., 2009). For this task, several images of *a group* of circles, with one (or more) distinctly larger than the others are displayed on a touch-screen mounted in the vehicle, and drivers are asked to distinguish the larger-sized circle(s) on the display and to touch/click on these. After touching all the larger circles required, they then click the "next" button to move onto the next scene in the task. The advantage of this type of *Circle Reference task* is that the levels of difficulty can be adjusted by varying the size and number of distracting or background circles – an example is shown in Figure 3.2, with three target circles highlighted:

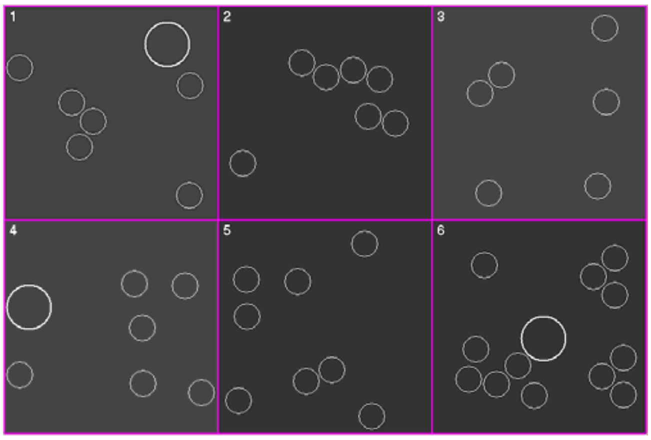


Figure 3.2 An example of Circle Reference Task

Again, three pilot runs (both desktop and on-road based) were conducted to test the suitability and applicability of the task design (and difficulty levels) on people's reaction time and effect on driving. The first pilot was conducted using a desktop computer and the second using a touch screen, while the third was on-road.

Participants for the first pilot were the same as those used for the auditory task described further above. The subjects were asked to react to the presentation of each scene by a left click of the mouse pointer on the targets circle(s). The results allowed three difficulty levels to be designed, based on a difference in subjects' reaction times (or the time duration between when a new scene appeared and the first correct click of the target circle), i.e. the reaction time increased when the task became more difficult. The average reaction time for task difficulty levels 1, 2, and 3 were 0.92s, 0.98s and 1.06s respectively, with a higher deviation as difficulty level increased (0.171, 0.176 and 0.236 respectively) - see Figure 3.3. The percentage of errors (or the number of times a target circle was not clicked, divided by the total number of clicks during one task) also increased with difficulty, at 3.36%, 10.12% and 13.16% for levels 1, 2 and 3 respectively. (Again, because of the small sample size, no statistical analysis was conducted.)

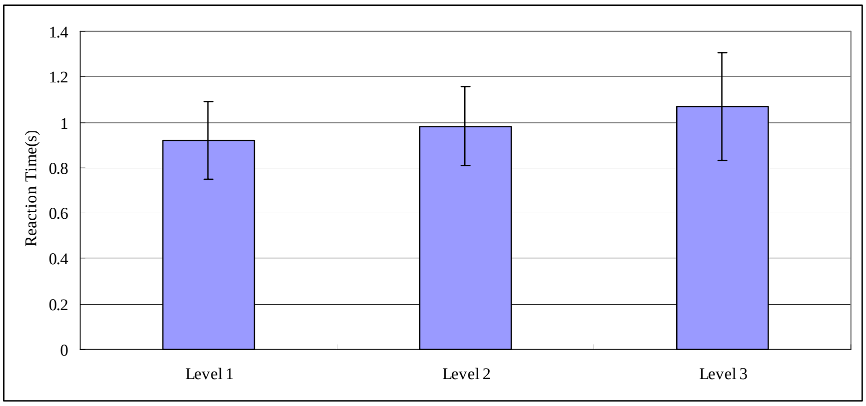


Figure 3.3 Reaction time on the visual tasks in the pilot run

Following the results and discussion raised in the first pilot, several changes were made to improve the visual task design, including:

1. an auditory cue was added every time a new scene appeared, to direct subjects' attention to the need for a new sub-task;
2. a time interval of one second was added between the end of one scene and the display of the next, to allow drivers time for adjustment.

Further improvements were made between the second and third pilots, which was again conducted by an experienced TRG technician and the experimenter. The touch screen was now divided into sections, and there was no more than one target in each section, to aid visual display (see Figure 3.2) and reduce the extra demand caused by any sun reflection on the screen. After the third pilot, the size of the circles were further enlarged to reduce the task demand for drivers, and numbers were added to each section of the display to offer a focus to drivers who found it difficult to manually click on the touch screen while driving. Figure 3.4 shows the final design of a visual task in operation.

Figure 3.4 Final design of the visual task – level 3

3.4.3 The Reference Task

In addition to the auditory and visual tasks, a reference task was designed before the experiment began to help assess how demanding these tasks were compared to a relatively simple IVIS task – turning on the radio. Recent studies have shown that the demand of turning on a radio has varied very little over the past 20 years, and turning on a radio (Bischoff, 2007) remains a robust benchmark against which to judge new in-vehicle systems. The reference task was again simulated using software developed for the Operations Simulation System, where drivers were asked to select a radio station using the touch-screen. Once a simulated interface (from the display used by a common IVIS system) is shown on touch screen, only three strokes are required to perform this reference task, and the subjects are asked to select “BBC 1” by clicking on the “Radio” button on the display, and then “FM Radio”, and then “BBC1”. A short radio clip was then played to confirm the correct completion of this task.

3.5 Selection of Participants

Participants were recruited through different channels to cover a wide spectrum of potential candidates, in order to randomise and reflect a broad demographic of people, backgrounds and occupations, and to reduce the impact of individual factors such as mental capacity and driving style. The recruitment channels included email, word-of-mouth, leafleting in Southampton, and company advertising. The recruitment and screening process considered the best balance of age, gender and driving experience (see below), using a driver “qualification” questionnaire (see Appendix A), which was completed by candidates beforehand. The questionnaire was also used to collect other characteristics information on drivers' annual mileage, and attitudes and experience of using IVIS (see Section 3.9 for further details).

After screening, 41 drivers (29 male and 12 female, aged from 21 to 65), who met the following requirements, were recruited to participate in the on-road experiment for this research. They all:

1. possessed a full driving licence, valid for used in the UK;
2. were over the age of 21 and less than 70 (i.e. deemed by the UK DVLA to be within "normal" driving capability);
3. had more than 3 years' driving experience and no major endorsements/penalty points on their license (maximum of 3 points);
4. had no medical condition or disability that precludes them from driving a motor vehicle on public roads; and
5. did not wear glasses while driving (for improved calibration and accuracy of eye monitoring).

All 41 recruited drivers were asked to participate in the on-road experiment, and a significant amount of time was invested to collect (and ensure the quality and completeness of) their IV, eye monitoring, secondary task performance, Subjective Rating and interview data during the experiment, as incomplete data*sets* cannot provide a complete picture of the complex relationships that exists between workload, driving performance and visual behaviour which is required for the purpose of this research. Despite the inherent complexity and difficulty in ensuring that driver datasets are complete, only 7 out of the 41 driver datasets were found to be incomplete due to technical failures (e.g. due to poor tracking and/or recording, or missing certain eye monitoring or IV data), and these were subsequently excluded from the analysis. Nevertheless, 34 fully complete, high-quality datasets from drivers were obtained, or 83% of the total driver data collected.

The driver were divided into different groups based on these 34 complete datasets, to investigate the impact of factors such as age and gender, and other driver characteristic, such as driving experience. The criteria for dividing these groups were decided by considering the balance of subject numbers in each group, and also the effect this has on the balance of other factor group (e.g. age versus gender). Several commonly used methods for dividing the groups were applied and tested, and the final decision made by using the lowest correlation of drivers' characteristics (of age, gender and experienece, etc) between groups and the best balance in each group. Details of the breakdown of each group, and the correlations between different driver characteristics, are discussed separately below.

3.5.1 Driver Demographics

Age

Three Age Groups were defined in this study:

- Age Group #1: 21 to 30 year olds (inclusive);
- Age Group #2: 31 to 40 year olds (inclusive); and
- Age Group #3: 41 to 60 years old (inclusive).

Gender

Male and female. The number of subjects in each age group and their gender distribution are listed in Table 3.1.

Table 3.1 Number of subjects and gender distribution in each Age Group

Age Group	Gender		
	Male	Female	Total:
#1	7	4	11
#2	8	4	12
#3	8	3	11
Total:	23	11	34

3.5.2 Other Driver Characteristics

Drivers were also divided into three Driving Experience Groups:

- Experience Group #1: those with less than 10 years' experience (inclusive);
- Experience Group #2: those with 11-20 years experience (inclusive); and
- Experience Group #3: those with over 21 years experience.

The number of subjects in each Experience Group and their gender distribution are listed in Table 3.2.

Table 3.2 Number of subjects and gender distribution in each Experience Group

	Gender		
Experience Group	Male	Female	Total:
#1	10	5	15
#2	8	2	10
#3	6	3	9
Total:	24	10	34

The information on drivers' annual mileage, experience and attitude towards using IVISs were collected using a qualification questionnaire where drivers were asked to rate these on a scale from 1 to 5. The numbers of drivers who ticked each value from 1 to 5 are listed in Table 3.3. In this table *"Expe_Audi"* stands for the Experience of using *Auditory* Systems (where points 1 to 5 denotes answers from *often* to *never* respectively); *"Atti_Audi"* stands for the *Attitude* towards using Auditory Systems while driving (points from *not distracting at all* to *very distracting*); *"Expe_Visu"* stands for Experience of using *Visual-Manual* Systems (from *often* to *never*); and *"Atti_Visu"* stands for the *Attitude* towards using Visual-Manual Systems while driving (from considered as *safe* to *unsafe*).

Table 3.3 Number of driver subjects - by answers to Driver Qualification Questionnaire

		Number of Subjects						Source
Characteristic:	Where *1* means:	*1*	*2*	*3*	*4*	*5*	Where *5* means:	
Mileage (miles/year)	Less than 5,000	9	13	9	1	2	More than 20,000	Q. B3
Expe_Audi	Often	0	0	1	12	21	Never	Q. C1
Atti_Audi	Not Distracting	0	8	12	11	3	Very Distracting	Q. C2
Expe_Visu	Often	8	9	8	9	0	Never	Q. C3
Atti_Visu	Safe	2	13	11	6	2	Unsafe	Q. C4

Note: the **Source** column shows the related question number from the Driver Qualification Questionnaire - see Appendix A for details.

In order to obtain sufficient number of subjects in each group, the subjects categorisation were reorganised based on the answers to the Driver Qualification Questionnaire:

- For "Mileage", all subjects were re-organise into three groups (i.e. Mileage Group #1, #2 and #3), where the small group number represents the least annual mileage. In this new categorisation, the original Mileage Group #1 and #2 (who choose 1 and 2 for QB3) were retained, and the groups answered 4 and 5 were combined and formed a new group: #3. Therefore, in the new category, drivers belonging to Mileage Group #1 (9 subjects) have less than 5 thousands miles of annual mileage, Group #2 (13 subjects) contained drivers who had an annual mileage between 5 and 10 thousand, while Group #3 (12 subjects) had more than 10 thousand mileage per year.
- Similarly, Expe_Audi was organised into two groups. Group #1 (13 subjects) included drivers who rated their experience as 3 and 4, i.e. *more* often; while Group #2 (21 subjects) included those who answered *never* (to the question C1).
- Two Expe_Visu (Experience in Manual Systems) Groups were also defined: Group #1 (17 subjects) represented those who use manual systems *more* often, or those who rated "1" and "2" for the Question C3; and Group #2 (17 subjects) represented those who use manual systems *less* often, including ones who rated "3" and "4" for Question C3.

- Drivers' attitude towards *auditory* systems was organised into three (Atti_Audi) groups based on Q. C2, with Atti_Audi Group #2 (12 subjects) represented *neutral*, or those who rated "3"; Group #1 (8 subjects) represented an attitude that considered auditory systems as *distracting*, comprising drivers who rated lower than "3"; and Group #3 (14 subjects) those who rated higher than "3", considering these systems as *not* distracting.
- Drivers' attitude towards *manual* systems was also organised into three (Atti_Visu) groups, based on Q. C4. Again, Atti_Visu Group #2 (11 subjects) represented *neutral*, or those who rated "3" for that question; Group #1 (15 subjects) those who thought manual systems are *unsafe*, included those who rated lower than "3"; and Group #3 (8 subjects) included those who rated higher than "3", or thought these systems are *safe*.

The numbers of subjects in each Driver Characteristic Group are illustrated in Table 3.4.

Table 3.4 Number of subjects in each driver characteristic group

		Number of Subjects			
Characteristic		1	2	3	
Mileage (miles/year)	Less than 5	9	13	12	More than 20
Expe_Audi	Less	13	21	N/A	More
Atti_Audi	Not Distracting	8	12	14	Distracting
Expe_Visu	Less	17	17	N/A	More
Atti_Visu	Safe	14	11	8	Unsafe

3.5.3 Correlation between Different Driver Characteristics

Correlation analysis was conducted to investigate the distributions of age, gender, driving experience and annual mileage across different groups. The results showed that the only significant correlation existed between age and driving experience (Ordinal by ordinal, Spearman Correlation, $r = 0.741$, $p < 0.001$). The effects of age and experience were therefore considered, compared and discussed together in the final analysis, and any correlations taken into account. Apart from age and experience, the group divisions had no bias on any other factors, for example, the age, experience and annual mileage factors in the two different gender groups were balanced, so that any differences between the two gender groups would be considered as being caused by gender difference only.

The correlations between the driver demographic factors (age and gender), driving experience, mileage and experience and attitude towards in-vehicle systems were also investigated. The results showed that, based on the current sample size, significant correlations existed in:

1) Age and Attitude towards Visual-Manual Systems: Spearman Correlation, r = 0.411, p = 0.016, i.e. the higher age group is more likely to think these systems are unsafe;
2) Experience on Auditory Systems and Experience on Visual-Manual Systems: Spearman Correlation, r = 0.373, p = 0.030, i.e. drivers who tended to use auditory systems more also used more manual operation systems;
3) Experience on Visual-Manual Systems and Attitude towards Visual-Manual Systems: Spearman Correlation, r = 0.485, p = 0.004, i.e., drivers who had more experience in using these systems were more likely to think their operation was safe.

However, these particular correlations had no significant impact on the findings of this experiment.

3.6 Apparatus

3.6.1 TRG Instrumented Vehicle (IV)

Experimental data were collected using the IV (see Figure 3.5), which consists of a Fiat Stilo motor car equipped with sensors to measure vehicle speed, location, distance to adjacent vehicles, driver's usage of controls, eye movements (via FaceLAB^{TM}), and other behavioural actions (McDonald & Brackston, 1997).

Figure 3.5 The Instrumented Vehicle

The sensors in the IV include (See Figure 3.6):

- a "VBOX" *digital Global Positioning System* (dGPS), to measure vehicle location (longitude, latitude and altitude), velocity and acceleration, travelling distance and vehicle yaw rate, with a measurement frequency of 100Hz, and an accuracy of 0.5m (for distance), 1.8m (positioning), 0.1 degree (direction), and about.0.1 km/h (speed);
- *front and rear-facing Radars*, to measure distance headway, relative speed and the acceleration of adjacent vehicles, with a frequency of 25Hz and an accuracy to 0.1m (headway), 0.36 m/s (relative speed) and 0.1 m/s^2 (in acceleration); their tracking range is between 5 and 150m, and the speed operation range is between 30 and 180 mph;
- *two infra-red Laser Scanners*, to measure the spacing with surrounding vehicles, with a frequency between 10 to 40 Hz, and an accuracy of 5cm; their tracking range is up to 70m and the maximum number of vehicles tracked is 20;
- *four video cameras*, to record pictures of the front, rear and in-cabin views of the vehicle, as well as the driver's foot movements; and
- *a lane detection system*, to measure lane width, lateral position, drift velocity and road curvature, with a frequency of 30Hz, and an accuracy of ±2cm (lane width), 2.0 to 5.5m (lane width range), ±2cm (lateral position), and ±0.12 degree (curvature).

The sensor readings are stored electronically into two on-board digital computers contained in the IV, which allows the data to be exported to another digital media for further processing and analysis.

Figure 3.6 Sensors in the Instrumented Vehicle

3.6.2 FaceLAB™ Eye Monitoring System

The IV is also equipped with a *FaceLAB*™ Eye Monitoring System, which is used to investigate drivers' visual behaviour. The System consists of two further cameras, which track the driver's eyes and head using infra-red imaging, and another on-board computer to receive and process the different types of eye and head movement data. For example, the System measures the driver's head pose, eye fixations, gaze angles, saccade and blink, with a frequency of 60Hz. FaceLAB™ includes dedicated software, which is used to extract and smooth the raw eye and head movement data, and convert it to binary or text format (i.e. a .txt file). The System also provides a model representation of the layout in the vehicle, which allows the operator (the experimenter) to define the areas of interest, e.g. windscreen, mirrors, dashboard, and general surroundings, to estimate which objects are being observed by the driver - see Figure 3.7. This figure shows the "objects layout" in the system display when the driver is facing the windscreen (shown as a red line initiated from the "head model") and looking at the right mirror (shown as a green line from the head).

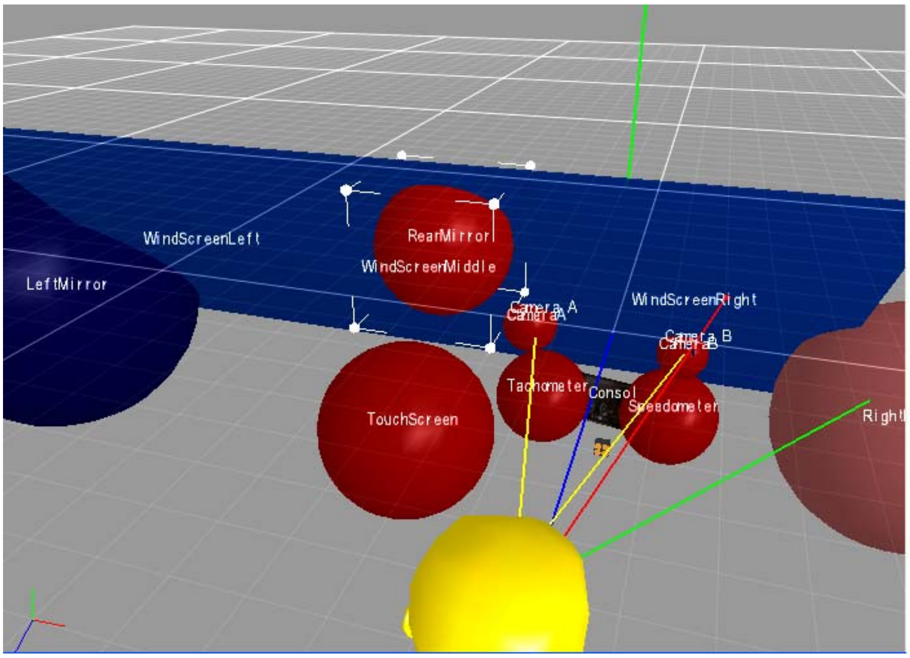


Figure 3.7 Pre-defined gaze objects/areas of interest in the FaceLAB™ system

For the purpose of this research, six areas of visual interest were defined, comprising left mirror, right mirror, rear mirror, speedometer, tachometer, and the touch screen. Each time the driver stares at one of these objects or areas, the camera records the focus of the gaze, and maps it onto the pre-defined area of interest in the System. According to the providers of the System (Seeing Machines, 2009), the tracking accuracy for the gaze angles and the driver's head-pose is ±0.5°, and live testing proved averagely this to still be within accurate tolerances (±1.0°).

3.6.3 The In-Vehicle "Operation Simulation" System

The IV has its own in-vehicle control interfaces, including a touch-screen and Fiat's exclusive (proprietary) voice-activation system. As these interfaces cannot be modified to meet the requirements of the research, an *Operation Simulation System* was developed specifically for drivers to perform the secondary in-vehicle tasks needed for the experiment and to help record the results. The in-vehicle Operation Simulation System ran on a laptop computer, linked to a separate touch-screen and speakers, which is connected to the IV. Figure 3.8 below shows the touch screen installed on the dashboard of the IV.

The System includes specialist "bespoke" software written for the purpose of this research, which simulates a set of IVIS visual and auditory tasks for drivers to perform. The software is written in the Visual Basic Software Development Language, which provides an in-vehicle "dialogue system" using a touch-screen and speakers, by which drivers can be presented with the secondary tasks to perform, and can respond to them. The System software and hardware required for the secondary tasks were developed and tested prior to the on-road experiment starting.

Figure 3.8 Inside of Instrumented Vehicle, showing location of touch screen used in experiment

The System provides two interfaces or sub-modules, which operate as the functions of the auditory or visual tasks. For each function, the experimenter clicks on one of the buttons on the software interface of the laptop, which activates the representative secondary task (via the speakers or touch-screen) to the driver. The software automatically records the times and other useful information of each *event* for subsequent analysis, for example when the task instruction is activated by the experimenter, or the time and the location of each stroke on the touch-screen.

Figure 3.9 shows the input and output flows for the System, including how drivers interacted with the System during the experiment, as they were asked to perform the secondary tasks.

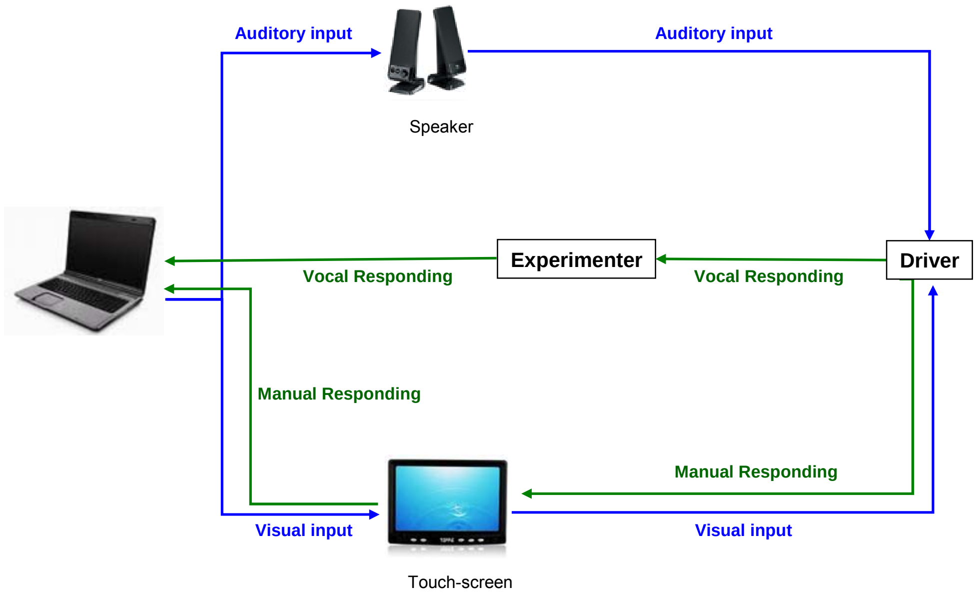


Figure 3.9 Inputs and output flows for the Operation Simulation System

In the *auditory* sub-module, the activation of each task sends an audio clip to the driver, which includes the information and actions required for that task - for example "Task level 3, the target sounds are *da, di* and *doe*, please start counting now" is followed by a sequence of different sounds which include several occurrences of *da*, *di* and *doe*. Once the audio clip has finished and the driver has responded vocally, the experimenter can either record their answer (and click "Next" to go to the next task), or can set the "Retry" button to replay the audio clip. In the *visual* sub-module, the System was developed to allow drivers to interact directly with the touch-screen, without intervention by the experimenter, apart from a prompt that it is a visual task. For each visual task, the driver received both the information (for example, images of circles of different sizes and locations) and performed the task (by selecting the "target circle") on the touch-screen. The timing and location of every touch was recorded by the System automatically.

3.6.4 Equipment Setup

The layout, positioning and function of the apparatus used in the experiment are shown in Figure 3.10.

Figure 3.10 The apparatus layout

The principal items of equipment are:

1) The on-board computer which manipulates the sensors and controls for data collection in the IV (only the display is shown);
2) The laptop used to run the FaceLAB^{TM} System, which is also used to record the eye movement data;
3) Two IR (infra-red) pods, attached to the FaceLAB^{TM} System, located in the dashboard;
4) Two cameras of the FaceLAB^{TM} system, used to observe drivers' eye and head movements, installed on the top of the dashboard (shown enlarged in the top right corner);
5) A second laptop, running the in-vehicle Operation Simulation System, which was operated by the experimenter (not shown) sitting in the rear right seat;
6) The touch screen used in the experiment mounted on the dashboard;
7) One of the two speakers, fixed to the back of the vehicle;
8) A monitor installed in the vehicle, which is used to observe the images of the four IV cameras (to confirm proper functioning of the IV sensors); and
9) A third laptop, which displays drivers' eye gaze, overlaid onto the scene ahead.

All the apparatus used for the experiment was tested and calibrated before *each* test run. (See Appendix F for details of the calibration procedure and the tests involved.)

3.7 Test Routes for the Driving Scenarios

As indicated earlier, two test routes representing two common driving scenarios (i.e. Car-Following and Free-Driving) were chosen for this study.

3.7.1 Test Route for Car-Following Scenario

In the Car-Following scenario, the primary driving task involves asking drivers to follow the leading vehicle at a constant headway, and drive in their normal way, without taking any additional risks. The driving route taken is illustrated in Figure 3.11.

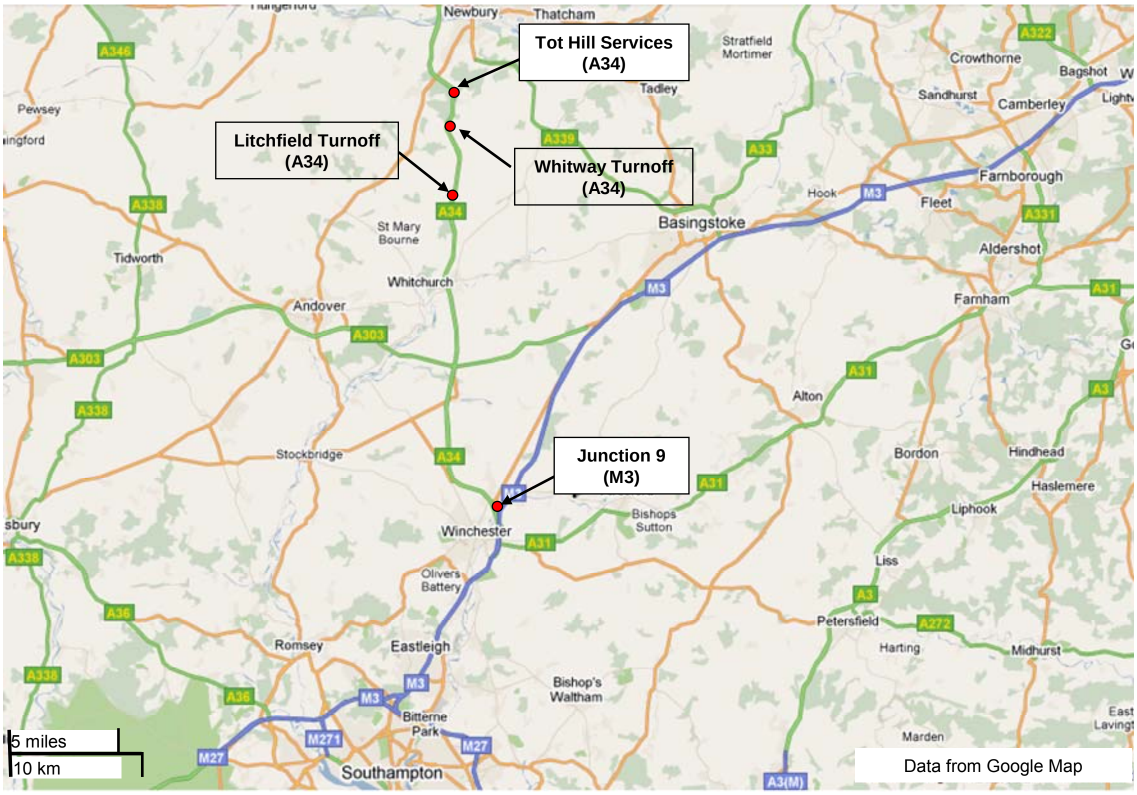


Figure 3.11 Test route for Car-Following Scenario on A34

The Car-Following Scenario covers a section of the A34 road in Hampshire. The A34 is a major route in England, which runs from Winchester in South Hampshire to the A6042 in Salford. It forms a major trunk route from Southampton to Oxford, Birmingham, The Potteries and Manchester. It is therefore practical for a Car-Following Scenario, especially lorry-following, and was found through previous TRG research to be a reliable, reproducible, and stable scenario for driving performance research (Zheng, McDonald & Pickering, 2008). The selected road section involves a dual carriageway throughout, with 2 lanes in each direction. For details of this road geometry, see the Design Manual for Roads and Bridges (Department for Transport, 2011).

The route consists of driving from a starting point from Junction 9 of the M3 Motorway in Winchester, and heading north along the A34 towards Newbury for approximately 21 to 26 miles, before returning back Southbound towards Winchester along the same road, at one of the three pre-determined “turning points”:

1) The Litchfield turnoff, also signed for Ardglen Industrial Estate (about 2-3 miles north of the Whitchurch junction);
2) The Beacon Hill turnoff (or about 2 to 3 miles further on). Also signposted for Whitway; and
3) The Tot Hill Services (at start of Newbury bypass). Also signed for Newbury (B4640).

The turning junction taken on each test depends on drivers’ progress in the secondary tasks, which are only started once the vehicle has passed beyond the national speed limit sign on the A34.

Drivers were instructed to stay in Lane 1 (or nearside lane), and preferably to follow a lorry at a comfortable speed (at c. 50 to 60mph), and to keep the headway as constant as possible. Where there was no lorry in sight, they were instructed to follow a car in Lane 1. The speed limit on this route is the national speed limit (For reference, the national speed limit for a dual carriageway is 70mph.).

3.7.2 Test Route for Free-Driving Scenario

In the Free-Driving Scenario, drivers were not required to follow any vehicle ahead. Instead, their primary driving task was to drive on a *consistent* speed. This Scenario used a section of the A33 in Hampshire, and the route was chosen because it is parallel to the M3 motorway, which takes most of the through traffic away, and therefore the test route was very quiet and ideal for the experiment. The route used is shown in Figure 3.12.

Figure 3.12 Test route for Free-Driving Scenario on A33

This route starts from the roundabout of the M3 at Junction 9, but then diverting from the A34 onto the A33, or the old primary road (pre-M3) from Winchester to Basingstoke. Drivers were asked to perform the secondary tasks after going over the old railway humpback bridge, on the climb up the hill away from Kings Worthy (just before a national speed limit sign). The speed limit on this road section is 60 mph, and drivers were asked to keep to a steady speed of 55 miles/hour (88 kph), as there was no requirement to follow another vehicle. The tasks were *only* performed when there was no vehicle ahead which could impact the drivers' speed choice and driving behaviour. |The secondary tasks were conducted on a single carriageway such as this for better experimental control, and two "laps" north and south along the A33 were required to fulfil the time required to perform all the secondary tasks. For details of the single carriageway road geometry, see Design Manual for Roads and Bridges (Department for Transport, 2011).

3.8 Experimental Surveys

Three surveys were used in this research for each driver:

- an initial driver qualification questionnaire;
- Subjective Dual-task Activation Ratings after each secondary task was performed; and
- A Post-Driving Interview after each driver trial.

These are discussed separately below.

3.8.1 Initial Driver Qualification Questionnaire

The purpose of the driver qualification questionnaire was to select the participants who meet the requirements for this research (as indicated in the "Participants" section above) and also to collect demographic information and drivers' experience and attitude towards using IVISs - see Appendix A.

After the experiment, the answers from the questionnaires were also analysed to see whether there were any correlations or trends which could help the subsequent data analysis for driving and/or secondary task performance and visual behaviour. The original hypothesis suggested that in-vehicle tasks should cause a higher task demand and therefore workload, which would lead to a deterioration in driving performance.

3.8.2 Subjective Dual-task Activation Ratings

Dual-task Activation Rating is a type of Subjective Rating, which is commonly used to measure drivers' workload. In general, *workload* is a multi-faceted phenomenon, which is related to the physiological states of stress and effort, and the subjective experiences of stress, the associated mental effort and time pressure, as well as objective measures of performance levels, and breakdown in performance. These various aspects of workload have led to separate and distinct means for assessing the workload, including physiological criteria, performance criteria, and subjective measures (Schvaneveldt et al., 1998). Between these three criteria, *Subjective Rating* and the method of Peripheral Detection Task (PDT) based on the residual amount of information processing resources are commonly used. Both have been proven to be a reliable estimation for workload (Jahn et al., 2005; Tsang & Velazquez, 1996). However, PDT will unavoidably increase drivers' workload, causing limitations in real-road driving applications. Given drivers' workload must be controlled to a limited level in on-road driving scenarios, it was decided that Dual-Task Activation Rating would be an optimal method in this study for estimating drivers' workload.

In the experiment, drivers were asked to rate the demand or workload experienced in performing each secondary task by comparing it to that of the Reference Task of turning on the radio, which has a designated value of 100. A similar ratings system has been used effectively and reliably in previous driving research (Angell et al., 2006), and such Subjective Rating is an optimal method for gaining feedback from drivers in experiments, because it is conduced immediately after the completion of each task. (Dual-task Activation Ratings means the ratings are given immediately after the event.) Each Rating is given vocally by the driver and recorded by the experimenter using a questionnaire-style sheet - see Appendix B.

3.8.3 Post-Driving Interview

An interview was completed at the end of each driving trial. This Post-Driving Interview was aimed at gathering information on drivers' objective opinions about the distraction they experienced in the different driving scenarios, and self-awareness of their driving behaviour while performing the secondary tasks. It contains three sections, which specifically investigate drivers' experience of performing the secondary tasks whilst driving, and separately any behaviour changes while performing either the auditory or visual tasks. The information gained from these interviews were then used to gain an improved understanding of drivers' experiences, by comparing their subjective opinion of any behaviour changes during the experiment with their actual behaviour as measured by the IV and the eye movement system. See Appendix C for an outline of the interview sheet used in the experiment.

3.9 Experimental Procedure

Ethical approval and risk assessments were applied during this research. They were submitted for evaluation on the 13th May, 2010 and approved by the Research Ethics Committee in the School of Civil Engineering and the Environment on 7th June, 2010 (Reference Number: CEE 2009/10-08), and submitted to the University of Southampton's Research Governance Office for approval on the same day, with approval given on 14th June, 2010 (Reference Number: 7304). The experiments were carried out in good weather conditions in August and September 2010, with each driver trial starting at either 9:30 in the morning or 14:00 in the afternoon to avoid peak hour traffic.

Participants were asked to fill in and sign a consent form prior to each trial, which were kept on record. A copy of the consent form is attached (Appendix D). It is stated explicitly on the consent form that participants are able to withdraw at any time they so wish, without penalty. For the experiment, the IV was driven along pre-defined sections of the A33 and A34 in Hampshire for approximately an hour in each Scenario (or 80 miles there and back). Drivers were asked to follow a leading vehicle on A34 keeping a constant headway, while on the A33 they were instructed to drive on a constant speed of 55 miles per hour, which allowed their performance to be assessed. For example, an increase in headway or deviation from the constant speed, can be taken as indications of deterioration in driver's performance while performing additional secondary (auditory or visual) tasks. See Appendix E for details of the instructions given to drivers. The order of the secondary task type and the order of the test routes/Scenarios were counterbalanced across subjects. The secondary tasks were only performed on each route shortly after the primary task (i.e. driving on the pre-designed test routes) has begun, and when an on-board safety observer (see below) deemed it appropriate to do so. To avoid the learning effects in performing the same types of secondary task (which can be similar in nature), the sequence by which the difficulty levels of the task to be performed by drivers was determined by counterbalancing across times of task repetition. Driving baseline data was also collected before, during and after the performance of the secondary tasks. After each trial run, the driver also attended an interview to give the experimenter details about their self-awareness of behaviour changes when performing the secondary tasks while driving. During the trail, each driver was accompanied by the experimenter and an independent safety observer (an experienced TRG technician), who was on the look out for traffic at all times. The driver participants were each paid a fee of £20 after the interview at the end of the trial.

The overall process for the trial is as follows (approximate durations are given in brackets):

1. Before each trial, the IV was checked to ensure the FaceLab^{TM} System and all other experiment equipment/instruments were in good and effective working order (30 minutes);
2. As each driver arrived, a check was made to confirm that he/she holds a valid driving licence, and possessed good eye sight (they had indicated this previously in the screening questionnaire). Participants were then given the experimental instructions (see Appendix E), and asked to read and sign the consent form (see Appendix D). (5 mins);
3. The requirements for the experiment were then explained to the driver. The operation of the secondary tasks was demonstrated and explained before any driving commenced. Drivers were given time to practice on the equipment for performing the seconday task (including the touch screen) until they were familar with the manipulation/actions required for each task (15 mins);
4. Before driving the IV, the FaceLabTM System was calibrated for each driver to account for their eye and facial dimensions and features (15 mins);
5. Drivers were then asked to drive the IV normally (without constraint) for half an hour along the M3 motorway, to allow them to become more accustomed to the vehicle (30 mins);
6. The primary task (e.g. following a leading vehicle) was then begun along the A33 and A34, where the IV and FaceLab^{TM} data collection began in earnest;
7. The safety observer would then check the road and traffic conditions (including the location of adjacent vehicles) to determine whether a particular secondary task could

begin; if conditions were suitable, the experimenter would give instructions for the driver to perform the secondary task through the in-vehicle Operation Simulation system, where the task type/difficulty level/interface sequence was determined beforehand (One hour for both the primary and secondary tasks);

8. Drivers were asked to rate the workload of performing each secondary task, using a Dual-Task Activation Rating compared to the baseline task of 100 for turning on the radio, e.g. a rating of 200 was judged to cause twice as much workload as turning on the radio (duration for this step included in that above);
9. Once the driving trial was finished, the driver was interviewed by the experimenter (15 mins);
10. After the trial, the data from the IV sensors and eye-monitoring system were downloaded for each driver for further analysis.

The total duration for each trial was therefore expected to take approximately two hours and 30 minutes, not including the initial IV/FaceLab™ System set-up and subsequent data downloading.

The names and contact details of the driver participants were kept in accordance with the Data Protection Act 1998. They were kept under the strictest confidence and were only used for the purpose of helping the objectives of this study, and no information from the data gathered will be published which might allow an individual to be identified. The Data were stored at a University PC, access to which requires the use of a password. Back up data were stored on a dedicated hard drive to avoid the possibility of data loss (e.g. due to PC failure), and stored inside a locked cabinet. The paper questionnaire responses were subsequently destroyed using University approved ‘data shredders’ following their transcription. Prior to this they were stored in a locked cabinet.

After the completion of the On-Road Experiment, the next Chapter will describe the data reduction and present database structure for the future analysis.

Chapter 4 Data Reduction and Database Structure

In Chapter 3, the on-road experiment to obtain data required for this research was described, i.e. the collection of subjective workload rating, secondary task and driving performance and eye movement data when performing secondary tasks, in real driving conditions. This chapter describes how these data were validated and processed, including some equipment calibration and data validation conducted before the experiment, to assure the quality of data used in the further analysis. The process of filtrating, data reduction and the processes of some complex data extraction are also introduced in this Chapter. As an outcome of this work, a structured database was established to extract the required information efficiently and accurately for further analysis.

4.1 Calibration and Data Validation

Equipment calibration and data validation both before and during the experiment are necessary to ensure that the experimental apparatus collects valid data which meet the requirements of this study. The calibration and validation process includes two major aspects:

- the IV data, i.e. data collected by all sensors installed on the IV; and,
- the eye movement data, i.e. data collected by the eye monitoring system, FaceLAB™.

The calibration and validation processes for such data were conducted both systematically before the experiment and randomly during the experiment. The following sections describe the data calibration and validation process before the experiment. The validation procedures for most parameters were conducted statically. The measurements in a dynamic situation, i.e. the real circumstance during the experiment, might be less accurate. However, the errors would occur when the vehicle is moving should be “white noise” caused by technical issues, therefore, will not cause bias for the results. Furthermore, the sample frequencies of the sensors are high enough (10hz the minimum) (see Chapter 3) to avoid errors caused by the relative changes in distances over a sampling interval.

4.1.1 Calibration and Validation for IV Data

The systematic calibration and validation process for IV data was carried out in July, 2010 before the experiment started, which involved a TRG technician and an experimenter. The data were collected on a selected site in the University of Southampton. The IV, a high speed camera, and a tape measure were used for the IV data collection and field measurements. In addition, another vehicle was also used as a moving object for IV radar data validation.

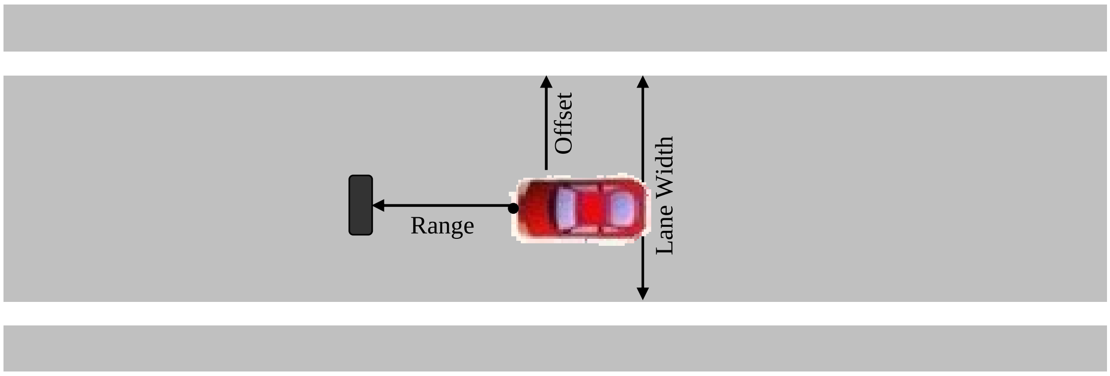


Figure 4.1 The diagram to validate the IV sensor data

The validation of Front Radar data was performed by measuring the distance between the IV and an object in front (another vehicle) because measured values of the distance between the object ahead and the IV (i.e. the *range* data) would start drifting towards zero, when the relative speed is low (or when the object is stationary). In the validation process, several distances of 15, 20, 30, 40, 50, 60, 70, and 100 meters were measured and recorded manually to compare against the range data collected from the Front Radar (as demonstrated in Figure 4.1).

The Lane Guidance sensor in the IV collects the lane width and the offset to the lane marking (i.e. lane position), see Figure 4.1. A similar process was applied to validate the sensor data by measuring the lane width and offset manually, and comparing the values between the IV data and the manual ones. Some comparison and validation results of such two sensors are demonstrated in Table 4.1, which suggest the accuracy and reliability of IV sensors of range and lane position.

Table 4.1 The validation results of the Front Radar and Lane Guidance Sensor

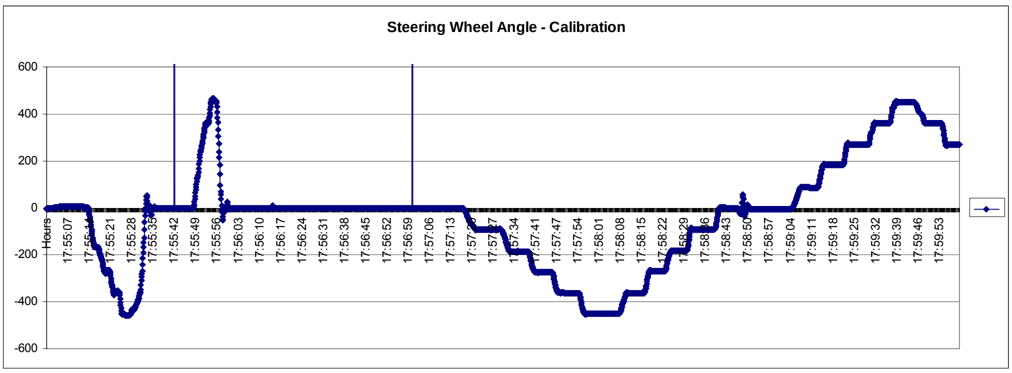
Sensor	Front Radar					Lane Guidance Sensor	
Parameter	Range #1 (m)	Range #2 (m)	Range #3 (m)	Range #4 (m)	Range #5 (m)	Lane Width (m)	Offset (m)
Sensor Measures	21.10	42.36	59.83	79.36	100.25	3.81	1.90
Manual Measures	20.98	42.28	60.11	79.42	99.78	3.79	1.89

The speed data from the IV CANBus were calibrated against the data collected by the VBox. The *Bland-Altman method* (Bland & Altman, 1986; Myles & Cui, 2007) was adopted to evaluate the consistency between these two speed measurements. The steering angle, brake pedal and pedal position were being constantly validated through similar processes with the validation of range and lane position data, i.e. comparing the IV data with manually collected and recorded ones. For example, the validation for steering wheel angles was conducted by comparing manual measures of the actual steering angles and the IV recorded data. Figure 4.2(a) demonstrates the process by which the steering turn angles were measured manually. Tapes were attached onto the steering wheel of the IV, crossing the wheel centre in straight lines vertically and horizontally respectively, as reference lines for the steering wheel movements. Therefore the four intersections on the steering wheel marked the angles of 0, 90, 270 and 360 degrees of steering wheel turns, respectively. When the steering wheel was turned, the time and steering angles were recorded both manually and by the IV. Figure 4.2(b) shows the reading from the IV when turning the steering wheel during the validating process when one small steering wheel turn was conducted at a time. The straight horizontal lines were observed between each movement (when the experimenter was recording the results). The measured values from two such different sources were compared against each other and the results are illustrated in Table 4.2, which validates the accuracy of the steering angle data. After several validation processes, the IV data were confirmed to be reliable and the accuracy enough for the purpose of this study.

Table 4.2 The validation results of the steering wheel angles

Time	17:58:00	17:58:20	17:59:00	17:59:20	17:59:40
IV Angles	-451.8	-269.4	-4.7	181.5	449.6
Manual Angles	-451.0	-270.0	-5.0	181.0	460.0

(a) A demonstration of measuring steering wheel turns manually



(b) The steering wheel angles recorded by the IV

Figure 4.2 The validation method for the steering wheel angles

4.1.2 Calibration and Validation for Eye Movement Data

The calibration and validation processes for eye movement data were conducted over a period of time from April 2010 to July 2010, both in-door and out-door, in order to factor in the effects of any change of lighting conditions which may occur during the experiment. The calibration and validation of eye-movement data include three categories of measurements:

- Gaze Angles;
- Parameters for *endogenous* eye movement: e.g. Blink and Saccade (rapid eye movment); and,
- Objects of Interest.

The procedure to validate the FaceLAB™ data was carried out with one subject (a TRG technician or one of the TRG students) sitting in the driver's seat, following instruction of fixating on certain objects, blink or transiting gazes from one object to another. Firstly, gaze angles were collected while asking the subject to fix their gaze on to several marked out-vehicle objects. The relative locations of fixed objects to the centre of the head of the subject were measured by a tape measure, and based on the relative location measures, the gaze angles (both *yaw* angle – horizontal gaze angle, and *pitch* angle – vertical gaze angle) from the visual central were calculated, as illustrated in Figure 4.3. Knowing the relative location of the object, i.e. certain x and y, the gaze angle θ is calculated by:

$$\tan\theta = x / y \qquad (4.1)$$

The gaze-angle data recorded by the FaceLAB™ system were then validated by comparing the FaceLAB dada against these *field* measurements (i.e. the calculated gaze angle θ).

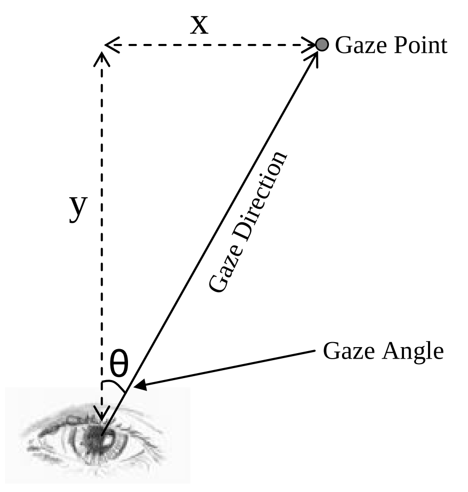


Figure 4.3 The method for manually measuring of eye's gaze angles

During this data collection process, the blink and saccade data were also obtained by subjects making blinks and saccades, and collected both automatically by FaceLAB^{TM} and by the high-speed camera mounted in IV capturing the driver's face image. The blink and saccade data were manually extracted from these video images, so that readings from the FaceLAB^{TM} can be validated against manually extracted data. Ten groups of gaze angles (evenly distributed in front of the driver's field of vision), along with blink and saccade data, were compared. Table 4.3 lists some results from this process, to demonstrate the validation process and accuracy of the FaceLAB^{TM} system (the eye monitoring system). The average errors of the pitch (vertical) angle and yaw (horizontal) angle were less than 1.5°. Figure 4.4 demonstrates an example for the error distributions of pitch and yaw angles at one point.

Table 4.3 The validation results of the gaze angle, blink rate, and saccade rate

Eye-Movement Parameters	Gaze Angle #1 (Degrees)	Gaze Angle #2 (Degrees)	Gaze Angle #3 (Degrees)	Blink Rate (Times per min.)	Saccade Rate (Times per min.)
FaceLAB^{TM}	5.0	10.0	20.0	24.75	12.67
Manual	4.8	10.8	21.1	25.00	13.00

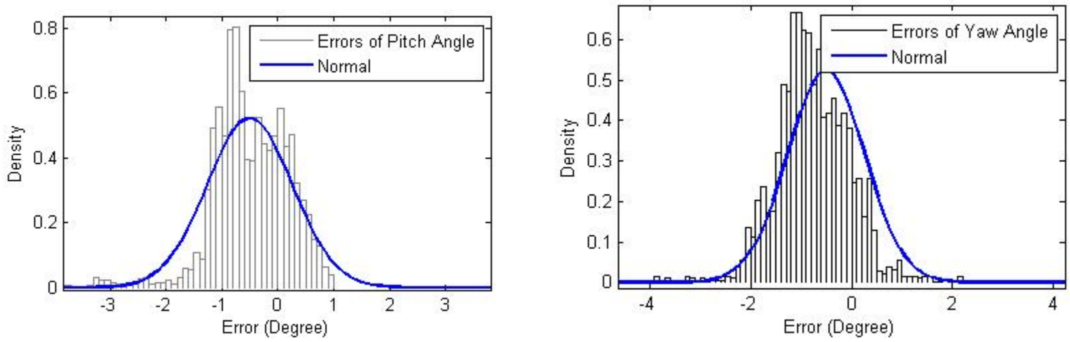


Figure 4.4 The error distributions of gaze angles

To determine the accuracy of gaze object data, eight areas of interest in the IV were defined in the FaceLAB^{TM} system. The subjects were asked to look at each of these eight objects for a duration of five seconds. In the validation process, subjects were instructed that they could feel free to move their head and hands while looking at the pre-defined objects, to simulate the situation people do in real driving. As data collected form the FaceLAB^{TM} were at a sample rate of 60 HZ, accuracy of the data was determined by the percentages of "gaze object" collected by FaceLAB^{TM} during the fixation on the corresponding gaze objects. Table 4.4 illustrates the accuracy of Objects of Interest from FaceLAB^{TM}.

The predefined objects/areas of interest are:

- Right Mirror;
- Speedometer;

- Tachometer;
- Touch Screen;
- Rear Mirror;
- Scene-View Camera (a small camera located in front of the subject, measuring gazes straight in front);
- Left Camera (the camera from FaceLAB^{TM} on the left);
- Right Camera (the camera from FaceLAB^{TM} on the right).

Table 4.4 The validation results for the Objects of Interest

Accuracy (Correctly detected in %)	Objects of Interest							
	Right Mirror	Speedometer	Tachometer	Touch Screen	Rear Mirror	Scene-view Camera	Left Camera	Right Camera
	98.7%	98.4%	99.1%	96.8%	97.8%	100.0%	95.7%	98.8%

4.2 Data Reduction

4.2.1 Re-Sampling

There was a need to conduct a data reduction into the same sample rate of 10 Hz, i.e., 10 readings per second, because not all sensors were sampled at the same frequency. For example, the Radar data collected produced a range output at a sample rate of 25 Hz. This output was therefore redundant under the fixed sampling rate (10 Hz), and needed to be re-sampled. Also, the output data from different IV sensors were saved into separate files, and each file only contained data from each sensor for five minutes duration. There was also a need to unitise all data formats, to combine or organise them into a structured database. It should be noted that the reduction of data from the Radar is much more complicated than the data from other sensors and the eye monitoring system. The following section describes the process of interpreting and organising this data.

4.2.2 Radar Data Reduction

Before the establishment of the database, some processing was needed for the data collected by the Front Radar, as the raw data could not provide straightforward leading vehicle information, as required for this research. This is because during an on-road experiment, the road environment and traffic conditions were much more complicated than in the IV-data validation process. The Front Radar equipped in the IV uses three beams to detect *targets* in front, and each beam can cover a maximum of four tracks (targets), and in total twelve tracks can be used to detect and record a target ahead. However, the Front Radar does *not* necessarily record one target in each of these twelve tracks separately, instead, a given target can be captured randomly by any track depending on the strength of the signal. In addition, the highly sensitive radar also tracks other on-road "noise", for example, lamp posts, metallic sign boards, and other passing or approaching vehicles, rather than the leading vehicle required for this study.

The objective of Radar data reduction is to obtain reliable *headway* data for the leading vehicle. Conventionally, the term *headway* refers to the distance from the front bumper to front bumper (or rear to rear) of one vehicle to the next one behind it (which includes the length of either leading or following vehicle), indicating the distance or time it will take for the trailing vehicle to cover that range. In this study, as this value was measured by radar device, for convenience, the gap between the rear bumper of the vehicle ahead and the front bumper of the IV (or *gap*) was used to describe Headway instead. It was found in the raw data that there always existed a strong Radar signal throughout in the Car-Following Scenario, and also, only eight out of the twelve tracks had significant responses while the signal strength from the other four tracks was very weak. This is because during the experiment drivers were instructed to follow the vehicle ahead, and the test roads were straight. Therefore only the eight tracks which contained the most information were selected for further headway data processing in order to reduce the background noise even further.

In the radar data, three parameters were used to describe each of the detected tracks: range (R_Range_Targ), relative speed (R_Velocity_Targ), and signal strength (Target_db), where the radar range represents the maximum distance at which the radar is effective in detecting objects. Table 4.5 provides a further description of the format, and examples for Front Radar data.

Table 4.5 The IV Front Radar output parameters

IV Sensors	Output parameters	Sample Values	Descriptions
PC Timer	Real Time	14:15:00	Computer time
Front Radar	ACC Speed	0	
	ACC Yaw Rate	0	
	R_Velocity_Targ_1	-8	Relative speed for Target 1
	R_Range_Targ_1	1688	Range (cm) for Target 1
	Azimuth_Targ_1	4926	Azimuth angle for Target 1
	Target_db_1	769	Signal strength for Target 1
	…	…	…
	R_Velocity_Targ_12	-12	Relative speed for Target 12
	R_Range_Targ_12	1888	Range (cm) for Target 12
	Azimuth_Targ_12	3912	Azimuth angle for Target 12
	Target_db_12	688	Signal strength for Target 12
	…	…	…

To process the remaining eight effective range readings, it was necessary to identify the targets of interest (in this case, the directly leading vehicle). There are two main factors that affect the target identification in the radar system: the *signal-to-clutter ratio* and the *resolution* of detection. The first factor is mainly controlled by the material property of the target: a metal target has a high signal-to-clutter ratio and a insulating target has a low signal-to-clutter ratio. The second factor is dominated by the radar geometry, i.e. the higher the beam width is, the lower the resolution. The front radar in the TRG IV has a vertical beam width of 3^o and a horizontal beam width of 8^o, therefore it has a better vertical resolution and lower horizontal resolution. The horizontal beam was divided into three overlapping angular resolution beams, each of which has a width of 3^o, providing a relation between the width of lane (*W*) and radar range (*R*) obtained as:

$$\tan(\theta/2)=W/2R \tag{4.2}$$

in which $\theta = 3^o$. Using the Taylor expansion, the above relation can be reduced as

$$R = 60\ W/\pi = 19.1\ W \tag{4.3}$$

Assuming the width of a lane to be $W = 3$ meters, the maximum range of the front radar will be R=57.3 metres. In other words, within the distance of 57.3 meters, the radar can detect vehicles in three straight lanes. Further than this distance, one beam may detect two vehicles or two beams may detect one vehicle. Figure 4.5 demonstrates the relation between beam angle, lane width and radar range.

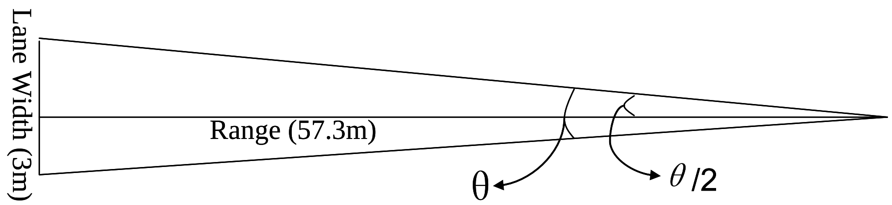


Figure 4.5 The relation between the width of lane and radar range

Therefore, even when there is no target ahead or targets are in the distance, the Front Radar may still capture strong signals reflected by guardrails, vehicles in other lanes or other on-road objects. As a consequence, the Front Radar would not be able to resolve targets properly without additional information. Given this complication, an efficient algorithm was designed to extract the target radar range from cluttered signals from the Radar. The following subsections describe the method, reasoning and logic of the algorithm, and a real-data example is also given.

4.2.2.1 Extraction Algorithm for Radar Range Signature

As discussed earlier, the required target (i.e. the leading vehicle) may be tracked by different radar beams and embedded in “noises” from clustered information due to complex on-road environment. Therefore in raw data, the required range information can be distributed in different tracks, and also, one track may contain range data for more than one target (e.g. other approaching vehicles). Therefore all range signatures from eight tracks should be combined together and stored in a single file. Figure 4.8 below shows an example of combined raw data from the eight tracks, where the x axis presents the time line and the y axis shows range value in each track. The cluttered range information from many different on-road objects can be observed.

The algorithm started with recognising the patters in *Range signature*. Range signature is the detailed range waveform (with respect to time) to identify the target. Different targets show different features in range signatures in terms of length, shape and fluctuation. From the example in the next subsection (see Figure 4.8), it was observed that the plot of the combined file is composed of numerous long or short curves and large amounts of isolated points, which contains both the required range signature and clustered noises. Some prior knowledge was applied to help the information extraction and the pattern recognition, for example, in the Car-Following Scenario, the leading vehicles are always in front of the radar, at a relatively constant distance. Therefore the range signature for the leading vehicle should be a continuous curve within certain amount of variances, during a relatively long period of time. While for vehicles approaching the IV from the opposite direction, the range signatures would be continuous but relatively short curves, with values of range reducing rapidly from maximum to zero. Furthermore, for over-taking vehicles, radar signature should appear as short curves with distance increasing from zero to a maximum. In addition, the signatures for stationary targets on the roadsides (e.g. lamp posts) would be very short curves or even scattered points. Therefore, the algorithm needs to recognise the pattern of the radar signature for the target as a long continuous curve, and exclude other irrelevant patterns. However, the long curve itself in a following scenario could also be interrupted by the leading vehicle changing lanes, which would therefore increase the effect of noise on correct range data extraction. Such interruptions should be *firstly* detected in the algorithm, before other filtering, smoothing, and interpolating methods were applied. These processes are discussed in detail in the following subsections.

4.2.2.1.1 Extraction of all curves from the combined file

To obtain the target range signature (e.g., the long and continuous curve) including potential “breakpoints”, all possible curves should first be extracted from the combined eight tracks of raw data. The procedure to extract each single curve (with *x* axis as time line and *y* axis as range values, see Figure 4.8) is:

1. Choose an arbitrary point as the beginning of a curve, and at this stage this point is also defined as the ending point for this curve.
2. Calculate the distances (*D*, i.e., differentials of *y* values) between the current ending point and *all* possible succeeding points (next point in terms of time line *x*). Determine the succeeding point by the minimum distance with the current ending point (*Dmin*), and update the new ending point by this succeeding one.
3. Repeat the above procedure until the minimum distance (*Dmin*) is larger than a threshold (*ThSpan*), i.e. *Dmin* > *ThSpan*, which suggests that the gap between the end of the curve and the closest potential succeeding point is too big to show a continuous line, and therefore the curve should be terminated. If the extracted curve is very short, e.g. smaller than a threshold (*ThCurLen*), it is regarded as a cluster or isolated points (for example caused by the radar reflection of a lamp post) and is not recorded or considered for further processing.

In the extraction process, attention needs to be paid on the *interruption points* due to intersection with another curve, because these interruptions can potentially lead the target curve in a wrong direction. To determine whether a point is an interruption point or not, a small *section* of data on the extracted curve is chosen (shown as a *solid* line in Figure 4.6) each time when a new succeeding point is added, and linear interpolation between the starting and ending points of this section is calculated (shown as *long dash* lines in Figure 4.6). If the distance of a point in the original curve to its corresponding interpolation line (shown as *dot line*) is larger than a threshold of *ThRaDiff*, this point is determined to be a interruption point; otherwise, if the distances between the original curve section and the interpolation line are all less than *ThRaDiff*, all points on this curve section are considered valid, as shown in Figure 4.7.

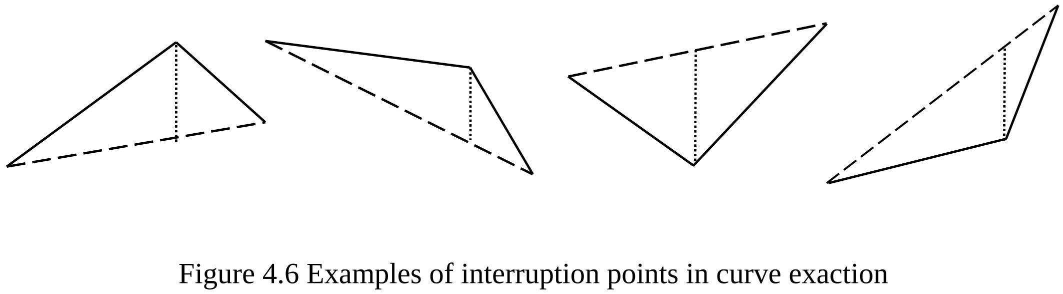


Figure 4.6 Examples of interruption points in curve exaction

Note: the long dashed line indicates the linear interpolation from the starting to the ending points of a curve section. In each of the four cases shown, the point at the *rooftop* (in solid line) is distant to the interpolation line, implying they are interruption points.

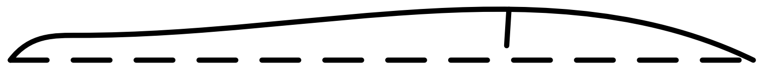


Figure 4.7 Normal target curve without interruption points

Note: the dashed line indicates the linear interpolation from the starting to the ending points and the solid line represents the original curve. In the normal case, the original curve is very close to the interpolation line implying that no interruption point exists.

4.2.2.1.2 Elimination of clustered curves

After the previous process had been carried out, all possible curves were extracted. However, most of them were “noises”, i.e. the radar signatures for other passing or approaching vehicles, rather than the target range for the leading vehicle, and therefore need to be distinguished and eliminated. The following steps were taken for this purpose:

- ***Generate a parameter matrix for all curves***. Each row of the matrix contains six parameters of each curve: the sequence number, length, positions of the starting and ending point, derivatives of the starting point and the ending point.
- ***Extract long curves***. Based on the curve lengths recorded in the parameter matrix, only long curves were extracted for further processing in the next step. When the lengths are smaller than a threshold, the curves are eliminated.
- ***Make sorts of the long curves***. Among the extracted long curves, the longest one is taken as a base for the range signature extracting of the leading vehicle. All other curves which appear on the same time line with the *longest* curve are eliminated because they do not constitute the target range signature, and therefore are considered as “noise” for the extraction in this study. The remaining curves are separated into two groups, i.e. the curves whose ending points are located on the left (on the *x* axis of Figure 4.9) to the starting point of the longest curve (i.e. left end of longest curve) as a backward group, and the long curves whose starting points are on the right to the ending point of the longest curve (i.e. right end of longest curve) as a forward group. The left and right here represent earlier or later in the time line respectively.
- ***Search in the backward direction***. Starting from the longest curve, make a search of effective curves in the backward group. Firstly, select the starting point of the longest curve as a reference point, and compare the derivative of this point and the derivatives of all ending points of the curves in the backward group. If the derivative of the reference point is equal or close to the derivative of one ending point in the backward group, then this curve is considered as part of the longest curve (i.e. target curve) with breakpoint; otherwise, it is an irrelevant curve to the leading vehicle and should be deleted. Update the longest curve by including the selected one on the left, and repeat the procedure until all curves in the backward group are searched.

- ***Search in the forward direction***. This is conducted by a similar procedure to the search in the backward direction, with swapping the starting point to the ending point of the longest curve, and searching for similar derivative of the left end of curves in the forward group.

4.2.2.1.3 Combination and interpolation of the recorded long curves

The extracted long curves contain the target range signature required, therefore such long curves were combined together to generate the final range signature. If any suspicious interfering curves still existed at this stage, they were screened and deleted manually using additional information from the on-road video images (collected by the IV simultaneously). The overlaps across different long curves were eliminated by *smooth average* technique. When a small gap exists between two curves (e.g. due to short signal missing), a linear interpolation was used to fill in the gap, and the video images were also used to validate this process.

4.2.2.2 An Extraction Example of the Target Range Signature

The algorithm proposed in Section 4.2.1 is powerful and efficient for the extraction of the required range signature from complicated clustered noises. An example of the radar data extraction is demonstrated in this section. A plot of the range signature using time as *x* axis and range values as *y* axis for the combined 8 tracks is shown in Figure 4.8, in which large amount of noise is observed. Using the first step (Section 4.2.1.1) of the algorithm, all continuous curves embedded in the clustered noises were extracted, as illustrated in Figure 4.9. It may be seen that the original data have been greatly reduced after this step, and the potentially valid, continuous curves were extracted successfully.

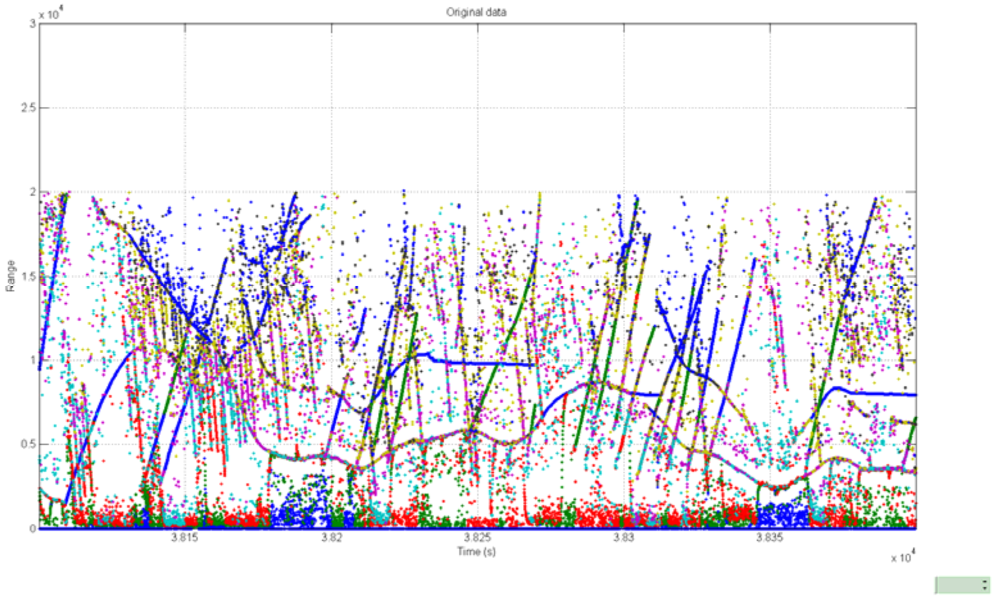


Figure 4.8 The original data for all range signatures of eight tracks

Although Figure 4.9 is much "cleaner" than Figure 4.8, some clustered curves are still blended in the target range signature, and further elimination using the second step of the algorithm (Section 4.2.1.2) is needed. After the second step, the signature is further distilled and only curves directly related to the leading vehicle range remain, as demonstrated in Figure 4.10, which provides the outline of the target signature. Finally, the third step (Section 4.2.1.3) is applied to combine the effective curves, screen out the suspicious signature curves, and interpolate the gaps between valid curves. As mentioned before, in this step, information from the on-road video images is used to validate the screening and interpolations. The final result is shown in Figure 4.11, which denotes the required range signature of the leading vehicle, showing the efficiency of the algorithm. The logic of the processes described above is very complex and computer programs were written by the author of this thesis to implement the various functions of this algorithm.

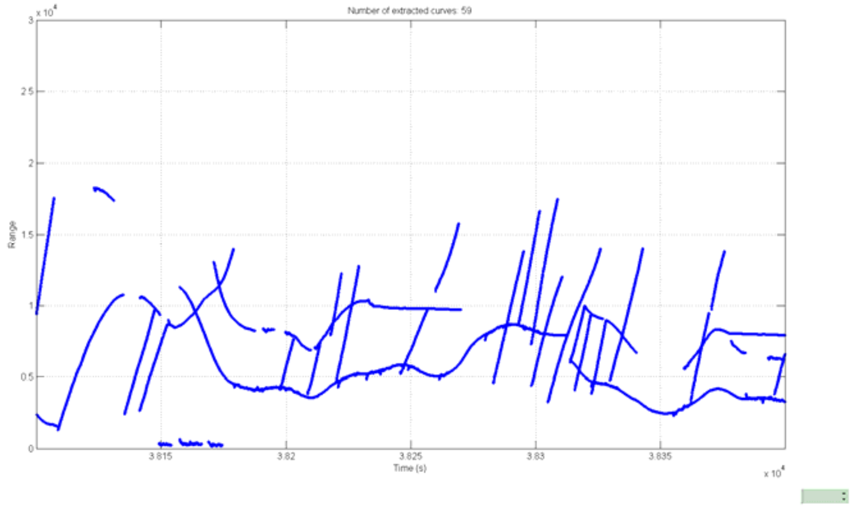


Figure 4.9 All continuous curves extracted from the original data

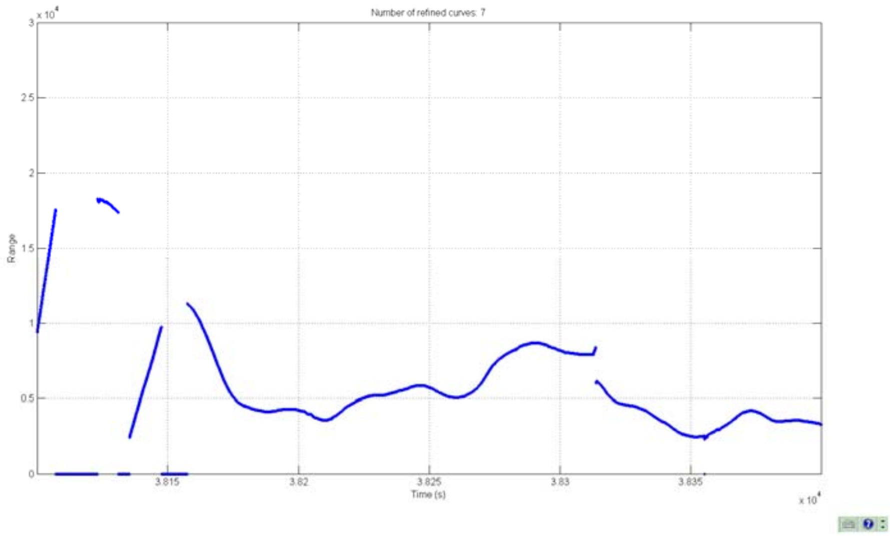


Figure 4.10 The remaining curves after eliminating the clustered curves

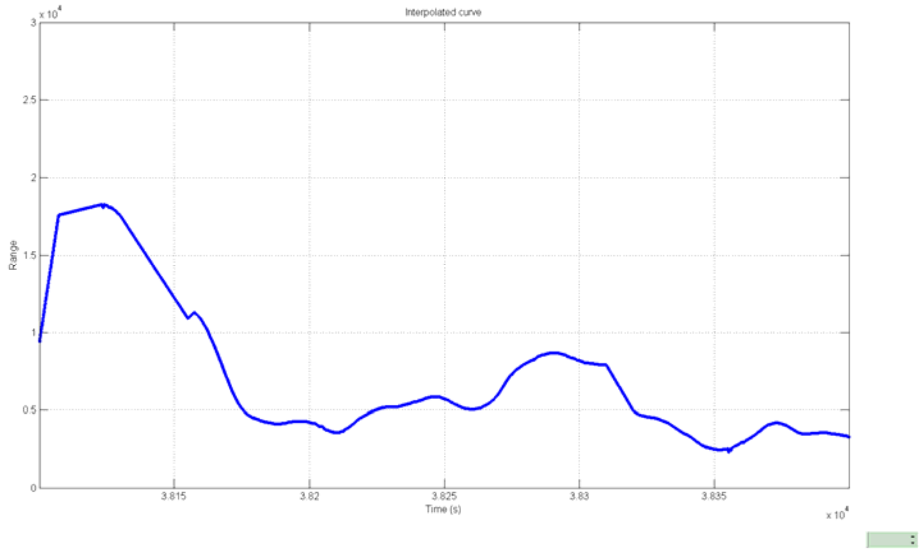


Figure 4.11 The final range signature after combination of effective curves and interpolation

4.3 Data Processing and Database Structure

Because this research is based on an empirical method where large amount of experimental data are essential, a good data structure is essential for the realisation of the research objectives. This section describes the process of constructing a universal database to evaluate drivers' workload, secondary task and driving performance, and eye movement behaviour. The database contains all data which were collected by different methods, from a large number of different files, in various formats and processed by specified methods. Three major databases were established: the workload and secondary task, the driving performance, and the eye movement database, because of the relative independence. Drivers' characteristics information was also important for further analysis; however for the completion of the On-Road Experiment (Chapter 3), it is described in the "Participant" Section.

4.3.1 Secondary Task Performance Database

As described in Chapter 3, drivers were asked to perform two types of secondary tasks in three different difficulty levels in two driving Scenarios. The structure of the secondary task performance database therefore contains all this information, as well as measurements for subjective rating on workload and secondary task performance. Drivers' auditory task performance was measured by the correct percentage of the answers (Correct Percentage), while performance of the visual tasks was measured by Reaction Time (RT), Standard Deviation of Reaction Time (SDRT), Error Percentage (ErP), and Error Distance (ErD). In this database, the performance measurements of secondary task were all calculated for each secondary task performed by each subject. Details of these measurements and reasoning are presented in

Chapter 5. In this Chapter, only the structure of the database is described, as illustrated in Table 4.6, in which the values contained in different columns are explained.

Table 4.6 The structure of the secondary task performance database

Column #	Variable	Remarks
1	Test_Num	The number of test run (from 1 to 41)
2	Driving_Scenario	The driving scenario (Car-Following or Free-Driving)
3	Task_Type	The task type (auditory task or visual task)
4	Task_Level	Task difficulty levels: 1, 2, or 3
5	Task_Ser_Num	The times of task repeating (1^{st}, 2^{nd} or 3^{rd})
6	Subjective_Rating	The subjective rating on workload
7	Correct_Percentage	The percentage of correct answers when performing auditory tasks (in %, valid for auditory tasks only)
8	RT	The reaction time (in second, valid for visual tasks only)
9	SDRT	The standard deviation of reaction time (valid for visual tasks only)
10	ErP	The percentage of errors (in %, valid for visual tasks only)
11	ErD	The distance of errors (in pixels, valid for visual tasks only)

4.3.2 Driving-Performance Database

Figure 4.12 depicts a flowchart of the procedure for establishing driving performance database. The experimental data from all 41 test runs were stored in different folders including all data from one test run, and each test run data contained hundreds of original data files named with different file extensions (each consists of five minutes data from a single sensor). These data were collected by various IV sensors and therefore all in different formats. In establishing the database, the first step was to process (including screen, format, filter, smooth and organise) the original data, re-sample them into a frequency of 10 Hz (the re-sample and smoothing methods used were from a previous work conducted in TRG, see Zheng, 2002), and organise them into a united matrix by the index of *time* (in Greenwich, synchronised by a Network Time Protocol, NTP server before each test run), for each test run (named Vehicle_x_10hz.txt, where x denotes the number of test run). Secondly, some calculations (for example acceleration and lane position) were conducted to generate a database for Driving Performance, again for each test run, named DP_x.txt (where x denotes the number of the test run). Finally, based on the task condition information which describes the driving conditions (e.g. baseline or task performing, driving scenario, leading and tailing vehicle types etc.) and starting and ending time of each task (in TaskPerformance.txt), and *following-episode* (each following episode refers to a duration when there is no leading *and* tailing vehicles changing lanes, see Section 4.3.2.2), data in DP_x.txt were divided into many segments. The statistical values of each segment were used to describe the driving performance during each segment, and this generated the final Driving Performance database, for further analysis (STDP.txt). These three steps are described in detail in this section.

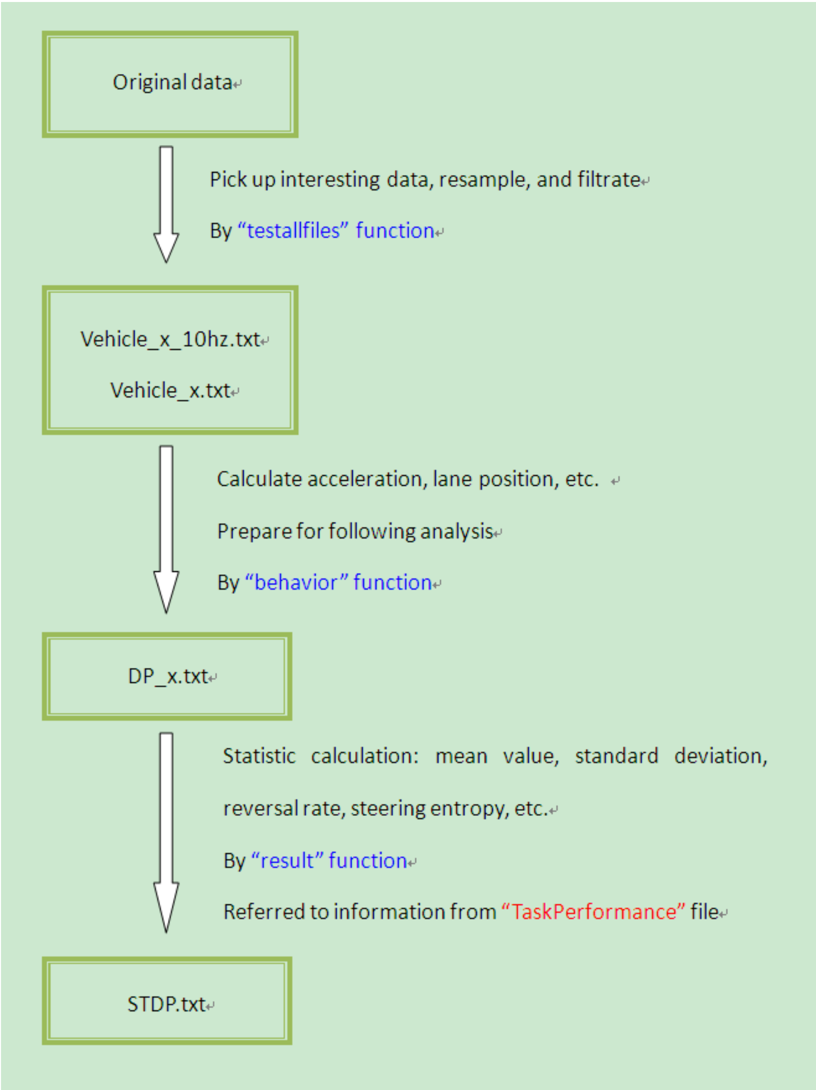


Figure 4.12 The flowchart of the establishment of driving performance database

4.3.2.1 Data processing: reading, sampling, and filtrating

A Matlab function was written to read all these files in different formats into 41 matrices, and each matrix contains all data for each test run (see Figure 4.13). Ten measurements were selected from the data matrices for further processing, including time, vehicle speed, brake force, gas pedal position, fuel consumption, brake intervention, steering wheel angle, event code, left lane offset, and right lane offset. These data were then sorted in time sequence and a time step of 0.1s was used for linear interpolation to *re-sample* the data into a sampling rate of 10Hz.

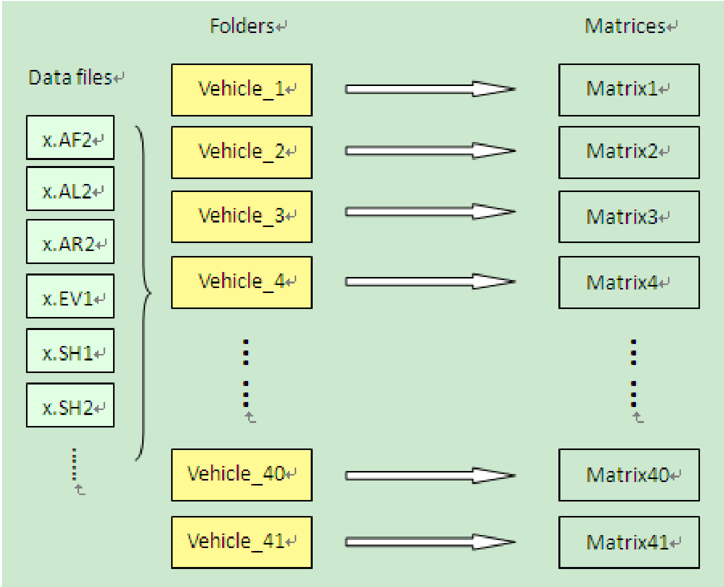


Figure 4.13 Extracting the data from all files in the 41 folders into matrices

Using the above procedure, multifarious data were read, sorted, and saved into 41 matrices into a uniform format, as illustrated in Table 4.7. The first column contains the time data in time sequence with 0.1 second step; while the following columns are all measurements data at the time corresponding to the first column. Suppose a test run started at time *t0*, the n^{th} values (i.e. the values for the time of *t0+0.1*n*) of measurement m: *Q (m, n)* is arranged in the $(n+1)^{th}$ raw and $(m+1)^{th}$ column. Some data in the matrices need to be filtrated, to smooth and excluded abnormal values due to technical errors. In the filtration process, human limitations of movement were considered to determine proper threshold. Figure 4.14 shows the plots of a data segment before and after being filtrated as a comparison. It can be observed that the high-frequency components (small oscillations) and errors were filtered out. This filter function was used to process the speed data, steering angle and lane offset data.

Table 4.7 The matrix formats of the selected data

Column1 time	Column2 Q1	Column3 Q2	Column 4 Q3	……	Column m+1 Qm
t0	Q (1, 0)	Q (2, 0)	Q (3, 0)	……	Q (m, 0)
t0+0.1*1	Q (1, 1)	Q (2, 1)	Q (3, 1)	……	Q (m, 1)
t0+0.1*2	Q (1, 2)	Q (2, 2)	Q (3, 2)	……	Q (m, 2)
……	……	……	……	……	……
t0+0.1*n	Q (1, n)	Q (2, n)	Q (3, n)	……	Q (m, n)

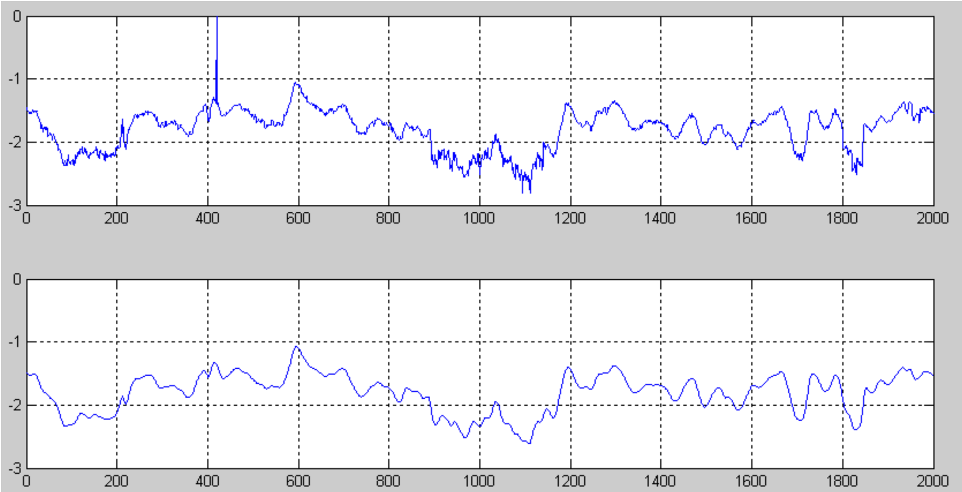


Figure 4.14 The comparison of data plots before and after being filtrated

After being re-sampled, filtrated and new measurements (i.e. lane position) being calculated, new matrices of the preliminary data processing were generated, which were saved as 41 text files respectively. The foregoing functions, i.e., reading, sampling, and filtrating, were all implemented by a MATLAB program. The processed radar data were also read, re-sampled, and then were combined into other vehicle data to form the "DP_x.txt" (where "x" represents the number of test runs from 1 to 41), and each of the DP.txt corresponds to all valid driving behaviour data from each test run.

4.3.2.2 The structure of driving performance database

Based on the constructed behaviour database, calculations were conducted to obtain the final driving performance database. As mentioned above, the driving performance data were calculated based on the duration of each task (including baseline) and *following-episode* information. The following-episode was to investigate the effects of leading (i.e. the vehicle immediately ahead the IV) and trailing vehicles (the ones following the IV) types on driving behaviour. Starting and ending times of a following-episode were extracted and recorded *manually* from the on-road video from the IV. Four vehicle types were considered in this research: passenger car (represented as vehicle type "1" in the following analysis), heavy goods vehicle (vehicle type "2"), van (type "3"), and no leading or following vehicle in sight (recorded as "0"). When going through all video images during each test run, whenever there was a leading or tailing vehicle change (including changing lanes), it was defined as a new *following-episode*, and the end of the current following-episode. The timing and the types of vehicles were then recorded. The starting and ending time of each following-episode were recorded from several seconds after the current episode begins and several seconds before it ends, in order to avoid the data fluctuation caused by the scenario change. From empirical practice, only the durations of following-episodes longer than five seconds were included in further analysis.

The driving performance database was firstly based on the duration of secondary task performance, i.e. each entry of data in STDP represents the driving performance during each task performance. However, if the following episode is not consistent during performing a task, the duration was considered as performing two or more different tasks. For example, performing the same secondary task with a leading change from a lorry to a van was considered as two different tasks. The driving performance data were therefore divided into smaller segments, with the starting and ending times for *both* task duration *and* following-episode taken into account, and the average values of the measurements were taken as driving performance during each segment. Therefore, in this database, each entry indicates the driving performance for each following-episode in the duration of each secondary task, named *"task-episode entry"*.

The reversal rate is a count of wheel reversal adjustments in a specified time period normalized to one minute. The discussion of the physical meaning and the importance to the road safety of this measurement is presented in Chapter 6. In this section, only the issues related to the data filtrating and reduction are discussed. There existed inevitable jitters in the steering wheel angles reading from the IV, which should be excluded. A function was designed to count the effective reversal times. Taking the steering behaviour in Test #1 from 15:04:57 to 15:06:00 as an example, the wheel angles data before and after processing are illustrated in Figure 4.15. The original data are shown in the top figure, demonstrating that the wheel turns frequently, the valid wheel adjustments, considered as intentionally manipulated by the driver, are marked as red circles. The bottom figure filtered out all the "noises", and only valid steering movements are shown, from which the two driving performance parameters: RR1.5 and RR3 (numbers of times drivers making steering adjustments larger than 1.5° and 3°per minute), were calculated. Steering Entropy is another parameter to evaluate steering behaviour, which is also calculated based on the filtrated data (for details about Steering Entropy, see Chapter 6).

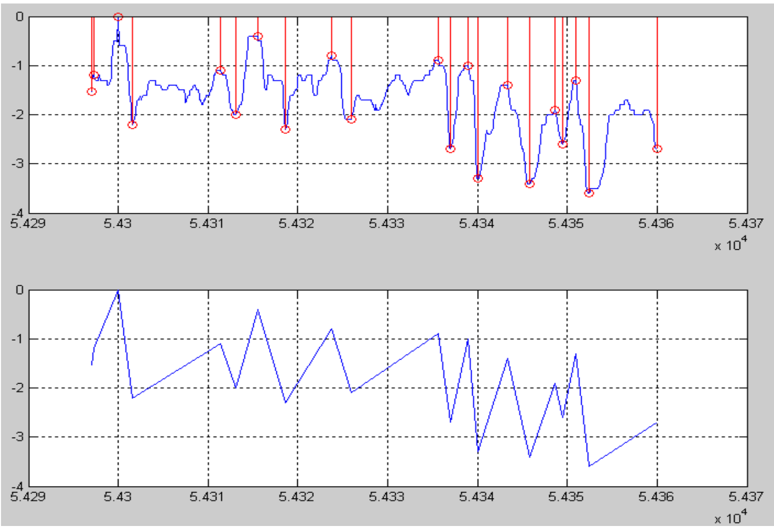


Figure 4.15 An example of filtrating for steering wheel angles measurement

The structure of the final driving performance database is illustrated in Table 4.8. Further explanation for these measurements, discussions on the reasoning why they were selected, and how they were applied to evaluate drivers' performance are presented in Chapter 5.

Table 4.8 The structure of the driving performance database

Column #	Name	Remarks
1	Test_Num	The number of test run (from 1 to 41)
2	Driving_Scenario	The driving Scenario (Car-Following or Free-Driving)
3	Task_Type	The task type (auditory task or visual task)
4	Task_Level	Levels of task difficulty, 1, 2, or 3
5	Task_Ser_Num	The times of task repeating (1^{st}, 2^{nd} or 3^{rd})
6	LVT	The leading vehicle type (0: no leading vehicle; 1: passenger car; 2: lorry; 3: van)
7	TVT	The tailing vehicle type (0: no tailing vehicle; 1: passenger car; 2: lorry; 3: van)
8	ST	The starting time for current record
9	ET	The ending time for current record
10	Duration	The duration for current record
11	MNSP	The mean speed
12	SDSP	The standard deviation of speed
13	SDLP	The standard deviation of lane position
14	RR1.5	The times of steering wheel turns larger than 1.5 degree per minute
15	RR3	The times of steering wheel turns larger than 3 degree per minute
16	Entropy	The steering entropy
17	MNHW	The mean headway
18	SDHW	The standard deviation of headway
19	MNTHW	The mean time headway
20	SDTHW	The standard deviation of time headway
21	MinHW	The minimum headway
22	MinTHW	The minimum time headway

All the functions and calculation described above were implemented by MATLAB programs, and the whole procedure of data organising is complex and time consuming, which took about three full months of tedious working.

4.3.3 Eye-Movement Database

The eye-movement data were collected by using the FaceLAB^{TM} eye monitoring system at a frequency of 60Hz. The original data from FaceLAB^{TM} were saved in four different files for each test run: timing, data from the *world model* (see Chapter 3), data for heading parameters, and for eye parameters. To analyze the eye-movement data efficiently, the four separate data files were combined for each test run. Similar approach of filtrating, smoothing, and organising with that for driving behaviour data were conducted for eye movement data. Fifteen eye movement measurements selected and calculated, along with task condition information, formed an eye movement database with 25 columns for all drivers. The structure of the database is described in Table 4.9. Further explanation for these measurements, discussions on the reasoning of why they were selected, calculations and some further validation of these measurements, and how they were applied to evaluate drivers' visual behaviour are presented in Chapter 7.

Table 4.9 The structure of the eye-movement database

Column #	Name	Remarks
1	Test_Num	The number of test run (from 1 to 41)
2	Driving_Scenario	The driving Scenario (Car-Following or Free-Driving)
3	Task_Type	The task type (auditory task or visual task)
4	Task_Level	1, 2, or 3
5	Task_Ser_Num	The times of task repeating (1, 2 or 3)
6	LVT	The leading vehicle type (0: no leading vehicle; 1: passenger car; 2: lorry; 3: van)
7	TVT	The tailing vehicle type (0: no leading vehicle; 1: passenger car; 2: lorry; 3: van)
8	ST	The starting time for current record
9	ET	The ending time for current record
10	Duration	The duration for current record
11	Nothing	The percentage of time spent on visual transitions in the current record (in %)
12	Panel	The percentage of time spent on the vehicle panel (Speedometer + Tachometer) in the current record (in %)
13	Mirrors	The percentage of time spent on the mirrors (Right mirror + Rear View Mirror) in the current record (in %)
14	Right_Mirror	The percentage of time spent on the right mirror in the current record (in %)
15	Rear_Mirror	The percentage of time spent on the rear-view mirror in the current record (in %)
16	TS	The percentage of time spent on the touch screen in the current record (in %)
17	Front	The percentage of time spent on looking ahead in the current record (in %)
18	SDGazeAll	The standard deviation on all gaze angle
19	SDGazeFront	The standard deviation on all gaze angle of looking ahead
20	Blink_Rate	The times of blink per minute in the current record
21	Blink_Duration	The average blink duration for the current record (in second)
22	Sac_Rate	The times of saccade per minute in the current record
23	Sac_Duration	The average saccade duration for the current record (in second)
24	Sac_Vol	The average saccade velocity for the current record (degree per second)
25	Sac_Amp	The average saccade amplitude for the current record (in degree)

After the database being established, the following Chapters 5, 6 and 7 will present the main analysis and results of this study.

Chapter 5 Subjective Rating and Secondary Task Performance

5.1 Introduction

Of the original 41 drivers who participated in the experiment, datasets from 7 test runs were excluded because incomplete data or poor eye tracking quality, as described in Chapter 3. All the results Chapters (5 to 8) are therefore based on the datasets collected from the remaining 34 test runs.

This Chapter presents drivers' **Subjective Ratings (SR)** on their workload and their **Secondary Task Performance (STP)** when performing the pre-designed in-vehicle auditory and visual tasks during the experiment, on the two pre-defined driving scenarios: Car-Following on A34 (**CF**) and Free-Driving on A33 (**FD**). The effects of driver characteristics including age, gender, driving experience, annual mileage, experience of, and attitude towards IVISs on the workload and secondary task performance were investigated using statistical models. The rationale of the statistical models used and a brief introduction of each model are also given in this Chapter. For the details of the secondary tasks and the driving scenarios, see Chapter 3. The analyses in this Chapter were conducted in the following four "*task and scenario"* types:

1. Auditory Task when Car-Following (ATCF);
2. Visual Task when Car-Following (VTCF);
3. Auditory Task when Free-Driving (ATFD); and
4. Visual Task when Free-Driving (VTFD).

The effects of the secondary tasks were compared between all three levels of task difficulty. As in the baseline driving (i.e., driving without performing any secondary tasks), drivers had no extra task apart from driving itself, the subjective rating and secondary task performance were not compared against baseline.

In this study, the subjective rating was collected by asking drivers to rate the demands of each task. A higher value of subjective rating indicates a higher experienced subjective workload. For more details, see Chapter 3. For auditory tasks, the task performance was measured by the **Correct Percentage** (**CrP** – the percentage of correct counting of *target sounds*); whilst for the visual tasks, the task performance was measured by **Reaction Time** (**RT** – the reaction time from the moment when each image presented, to the time when a driver responses to it),

Standard Deviation of Reaction Time (**SDRT**), **Error Percentage** (**ErP** – the percentage of clicks outside the 'target circle', i.e. percentage of the errors), and clicking **Error Distance** (**ErD** – the distance between the location of an error click to the closest 'target circle', in pixel, i.e., the 'extent' of mistakes).

5.2 Statistical Models

The data were analysed using the statistical methods. The statistical models used in this study, for example, ANOVA, Post Hoc Analysis models (e.g. Student-Newman-Keuls test And Tukey test), Non-parameter analysis, and Robust ANOVA models are introduced in this Section.

ANOVA is a general technique to test the hypothesis that the means between two or between more groups are equal by comparing the variances in the data sets using F-ratio, under the assumption that the sampled populations are normally distributed. The F-ratio is a measure of how different the means are, relative to the variability within each sample. The larger this value is, the greater the likelihood that the differences between the means are due to something (namely the real effect) other than chance alone. The F-ratio is calculated using mean square between different groups (MS_{betwee}) divided by the mean square that is due to the chance (MS_{within}):

$$F = \frac{MS_{between}}{MS_{within}},$$

where

$$MS_{between} = \frac{SS_{between}}{df_{between}} = \frac{SS_{between}}{k-1}$$

$$MS_{within} = \frac{SS_{within}}{df_{within}} = \frac{SS_{within}}{N-k},$$

in which,

- k = the number of different groups;
- *N = total number of all the values combined;*
- *SSbetween is the explained variation, i.e. sum of squares representing variation between the different* samples, which is calculated by

$$SS_{between} = \sum\left[\frac{(S_j)^2}{n_j}\right] - \frac{\left(\sum S_j\right)^2}{N};$$

- *SSwithin* is the unexplained variation, i.e. the sum of squares representing variation within samples due to chance, which is calculated by

$$SS_{within}=SS_{total} - SS_{between}$$

where

$$SS_{total} = \sum x^2 - \frac{\left(\sum x\right)^2}{N},$$

in which, n_j = the size of the jth group;

s_j = the sum of the values in the jth group;

x = one value: $\sum x = \sum s_j$,

As ANOVA uses variance to estimate whether the means are equal, there are several assumptions in ANOVA:

- the data points must be independent from each other;
- the population distributions must be normal;
- the variances of samples are not different;
- all individuals must be selected at random from the population;
- all individuals must have equal chance of being selected;
- sample sizes should be as equal as possible although some differences are allowed.

Before conducting each of the tests, the assumption of the ANOVA test was examined. Exploring Data Analysis (EDA) was conducted first. The normality of the underlying distribution was tested by comparing the distribution against a quartile plot and a probability plot of the normal distribution, and Leven's Test was conducted to scrutinise the "equal variances" assumption. The ANOVA test will not be affected much even if the background distribution is slightly skewed, unless the sample sizes are small. It is not the case in this study, because a relatively large database has been established. Therefore, when possible, an ANOVA test is applied. If this assumption of the ANOVA test was violated, the resulting F-test was invalid. In this case, an alternative test was conducted.

The Post Hoc Tests (multiple comparison tests) have been conducted when a difference was shown between more than two groups, to identify where exactly the effect existed. The specific model used was decided in the context of which was the most suitable. Hilton and Armstrong gave a detailed solution for different instances (Hilton & Armstrong, 2006). There are several parametric "multiple comparison tests", all of which carry out the same function. The two most common are the Tukey test and the Student-Newman-Keuls test. The Tukey test is reputed to be the most powerful (Zar, 1984) and would be the safest to use. However, The Tukey Test relies on the homogeneity of variance. Therefore, when this assumption is violated, the Games-Howell Test was used as an alternative, because it does not require “equal variance”. When the nature of a test is to compare the performance of several groups against a control group, not only the Tukey and SNK tests were applied, but also the Dunnett’s Test (Dunnett, 1955) was used to compare the differences between the baseline and levels of task difficulties. Because this test provides a powerful test of significance while controlling Type I error to a reasonable level (5%) when comparisons are made only between treatment groups against a control group (Klockers & Sax, 1986).

Outliers are generally considered as data points that are far outside the norm of a variable or population (Jarrell, 1994). The presence of outliers can lead to inflated error rates and substantial distortions of parameter and statistical estimates (Osborne & A., 2004). Therefore, outliers have been checked and treated with care before conducting the data analysis. Using the simple rule of thumb (Wainer, 1976), the outliers are defined as such points locating near three standard deviations from the mean, which may have a disproportionately strong influence on the parameter estimates in this study.

The unequal variance is another issue when using ANOVA. The unequal variance increases the chance of reporting a significant difference when there is none (Type I Error), especially when the sample sizes are not equal. In this case (i.e., unequal variance and unequal sample size appear at same time), because a non-parametric test such as Kruskal-Wallis still assumes that the population variances are comparable, even non-parametric tests would not be suitable. Therefore, the “corrected version of ANOVA” – Robust ANOVA tests, e.g. the Welch Test and Brown-Forsythe Test (BF Test) are employed when the Levene’s Test suggests the violation of “equal variances”. Some more details of Robust ANOVA tests are provided below.

The Brown–Forsythe test is a statistical test for the equality of group variances by performing an ANOVA on a transformation of the response variable. The transformed response variable is constructed to measure the spread in each group. Let

$$z_{ij} = \left| x_{ij} - \tilde{x}_j \right|$$

where $\tilde{x}_j$ is the median of group *j*. In fact, the Brown-Forsyth test is the F statistics resulting from an ordinary one-way analysis of variance on the absolute deviations from the median:

$$F = \frac{N-k}{k-1} \frac{\sum_{j=1}^{k} n_j (z_{.j} - z_{..})^2}{\sum_{j=1}^{k} \sum_{i=1}^{n_j} (z_{ij} - z_{.j})^2}$$

in which k is the number of groups, n_j is the number of observations in group j, and N is the total number of observations.

For Welch Test, when assuming normality, the null distribution of W can be approximated by the F-distribution. The difference between the F-statistic and the W-statistic is that the denominator of the F-statistic is based on the pooled sample variance, whereas, in the denominator of the W-statistic, the variances of the samples are considered separately (Buning, 1997). For a detailed introduction to Robust ANOVA, see Brown (Brown & Forsythe, 1974) and Welch (Welch, 1951). The Robust tests were shown to be robust against heterogeneous variances (Brown & Forsythe, 1974). Using Monte Carlo simulations, researchers confirmed that Robust Tests performed the best when the normality and equal variance assumption for ANOVA were violated, however, in the case when the variances were equal, non-parametric tests were more preferable (Buning, 1997; Roth, 1983). Therefore, for the non-normal but equal variance data, a non-parameter test was conducted.

Non-parameter tests (Mann-Whitney U Test and Kruskal-Wallis Test) are employed when the distributions are skewed, but still kept an equal variance between groups. Also, the ANOVA test can be seriously affected if the sample sizes are unbalanced, while as long as the variances are equal, non-parameter tests are robust for unequal sample sizes. When this occurs, a non-parametric test was conducted. The Mann-Whitney U test is often viewed as the nonparametric equivalent of Student's t-test. Like the parametric Student's t-test, the non-parametric Mann-Whitney U test:

-- is used to determine if a difference exists between two "groups," however you define "group";

-- is ideally dependent on the random selection of subjects into their respective group.

The major difference between the Mann-Whitney U Test and Student's t-Test involves the concept of normal distribution. Another non-parameter test employed in this study is Kruskal-Wallis Test, which is a generalised version of Mann-Whitney U Test. It makes a comparison between the medians of two or more samples to determine if the samples have come from different populations. The Post-Hoc for Kruskal-Wallis was calculated by hand using paired comparison with the Bonferron-corrected method.

The assumptions of non-parameter tests are less strict compared to ANOVA:

- the data points must be independent from each other;
- the distributions do not have to be normal and the variances do not have to be equal;
- you should ideally have more than five data points per sample;
- all individuals must be selected at random from the population;
- all individuals must have equal chance of being selected; and
- sample sizes should be as equal as possible, although some differences are allowed.

Depending on the distribution and other factors (e.g. existing of outliers), when necessary, several tests were conducted for the same research questions, and the results were compared side by side to gain the most convincing conclusion. Also, for some parameters when the distribution is heavily skewed and the compared variances are unequal, data transformation is conducted, in order to form a good shape of the distribution. These results will be reported when different from statistical results from the original parameter. If it is not specified, the level of $p < 0.05$ is considered as significant in this thesis.

As summarised by Yu (Yu, 2011) based on the results from Skovlund and Fenstad (Skovlund and Fenstad, 2001 cited in Yu, 2011) who compared the Type I error of the T-Test, Whitney-Mann-Wilcoxon Test (WMW), and Welch Test with variations of three variables: variances (equal, unequal), distributions (normal, heavy-tailed, skewed), and sample sizes (equal, unequal), the most suitable statistical tests recommended for each scenario are listed in the following table:

Table 5.1 Recommended Statistical Tests for Different Dataset (Yu, 2011)

Variances	**Distributions**	**Sample sizes**	**T-Test**	**WMW Test**	**Welch Test**
Equal	Normal	Equal	*	+	+
		Unequal	*	+	+
	Heavy tailed	Equal	+	*	+
		Unequal	+	*	+
	Skewed	Equal	_	*	_
		Unequal	_	*	_
Unequal	Normal	Equal	+	_	*
		Unequal	_	_	*
	Heavy tailed	Equal	+	_	+
		Unequal	_	_	+
	Skewed	Equal	_	_	_
		Unequal	_	_	_
Symbols: * = method of choice, + = acceptable, - = not acceptable					

5.3 Pre-Analysis: Learning Effect

As all three levels of the auditory and visual tasks were conducted three times on each test route, there was a need to check the learning effects from the repetition of secondary tasks on the SR and STP. The analysis was to investigate whether the workload changed and/or the secondary task performance was “improved” or not because of the repetition. If there was a significant difference caused by the repetition of the tasks, either the learning effect would be taken into account as an impact factor, or some data will be excluded to screen this factor out. The analysis of learning effect was performed in the following four scenarios.

5.3.1 Learning Effects on SR and STP in ATCF

SR and CrP data were skewed and long-tailed, and outliers were found in both datasets in EDA. Long-tail refers that a larger share of population rests within the tail of a probability distribution than observed under a Normal distribution. The Levene’s Test confirmed the homogeneity of subjective rating and correct rate for repeating of tasks, i.e. distributions of the data from the three times of performing the auditory tasks were comparable. The non-parameter tests were therefore chosen as a more reliable and powerful method to study the learning effect on the current database, without further data trimming and transforming. Table 5.2 shows the mean values of subjective rating and correct percentage in each task level when the same task was repeated.

Table 5.2 The Learning Effects on SR and STP for the case of ATCF

Task Level	1			2			3		
Times of Repeating	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
SR	86.56	93.94	95.13	132.97	123.41	121.47	178.24	191.75	167.38
CrP	85.29%	84.34%	90.03%	76.47%	82.35%	83.82%	55.05%	51.04%	66.67%

The Kruskal Wallis Test was used to test the null hypothesis H0: there were no changes on subjective rating and task performance caused by repetition of tasks. The results showed that the repeated performing of auditory tasks had no significant impact on the subjective rating and correct rate in the Car-Following Scenario ($p > 0.05$). Hence the learning effects on SR and CrP for the case of ATCF were not considered in the further analysis.

5.3.2 Learning Effects on SR and STP in VTCF

All data of the parameters for visual tasks performance in Car-Following were skewed. The homogeneity in variance test confirmed the equal variances from three datasets. The non-parameter test – Kruskal Wallis Test - was therefore more suitable for this analysis. No statistically significant learning effect was found for all chosen parameters (p > 0.05). The mean value of SR, RT, SDRT, ErP and ErD in each task level are illustrated in Table 5.3.

Table 5.3 The Learning Effects on SR and STP for the Case of VTCF

Task Level	1			2			3		
Repeating of Tasks	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
SR	103.26	86.76	85.50	137.18	118.35	121.82	171.21	160.24	147.94
RT	2.43	2.26	2.14	2.49	2.45	2.42	2.95	2.76	2.75
SDRT	1.28	1.10	0.85	0.68	0.68	0.87	1.02	0.84	0.87
ErP	16.17%	12.68%	9.87%	12.00%	11.30%	12.18%	11.82%	11.03%	12.07%
ErD	8.65	10.22	7.63	7.81	8.84	13.02	11.58	6.87	7.77

5.3.3 Learning Effects on SR and STP in ATFD

Similar analysis was conducted for the learning effect of the auditory tasks in Free-Driving, and no significant learning effect was found either (p > 0.05), see Table 5.4.

Table 5.4 The Learning Effects on SR and STP for the Case of ATFD

Task Level	1			2			3		
Repeating of Tasks	1	2	3	1	2	3	1	2	3
SR	82.03	90.06	90.97	132.12	121.32	121.94	183.84	171.57	172.16
CrP	97.06%	93.94%	87.50%	79.41%	85.29%	72.58%	61.30%	65.56%	74.20%

5.3.4 Learning Effects on SR and STP in VTFD

In the case of visual tasks in the Free-Driving Scenario, SDRT and ErD were skewed, but all parameters showed homogeneity between three times of task repeating. For the same reasons, the non-parameter test was employed here. The mean values of SR, RT, SDRT, ErP and ErDis in each task level are illustrated in Table 5.5. For all parameters, the learning effect was not significant, apart from a marginal improved accuracy in level 1. This improvement was however minor, less than 1mm (i.e. ErD from 19.42 to 13.14 pixels on a 5 inch touch screen, p = 0.068).

Table 5.5 The Learning Effects on SR and STP for the Case of VTFD

Task Level	1			2			3		
Repeating of Tasks	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
SR	105.26	92.36	93.35	140.44	123.74	123.34	173.88	173.32	155.59
RT	2.89	2.53	2.49	2.87	2.78	2.44	3.03	3.14	3.11
SDRT	1.44	1.02	1.30	1.15	0.96	0.90	0.94	1.03	1.02
ErP (%)	18.48	14.41	18.67	15.04	15.57	14.51	15.97	14.16	16.74
ErD	19.42	15.04	13.14	15.50	12.13	10.69	9.14	12.59	8.68

5.3.5 Summary of Learning Effects

In conclusion, the repetition of performing the secondary tasks had very little effect on the subjective rating and secondary task performance in all cases. Some improvements in task accuracy were observed (for example, increasing correct percentage in ATCF, and decreasing error distance in VTFD), but none of these effects was statistically significant. This may due to well randomisation of the scenario and tasks in the experimental design. It was therefore believed the learning was insignificant, especially when comparing to the effects caused by the secondary tasks, and it was decided to exclude learning effect as an impact factor in the main analysis.

5.4 Subjective Rating (SR)

This Section reports the impact of task difficulty along with the factors of drivers' characteristics on drivers' subjective rating, which indicates their workload. As discussed before in Chapter 3, a Dual-Task Activation Rating was used to collect the subjective rating data, where a higher value represents a higher experienced workload.

5.4.1 Outliers

It was observed that there were some extreme values in the subjective ratings. The presence of such extreme values (i.e. outliers) can lead to substantial distortions of parameters and statistical estimates when using either the parametric or nonparametric tests (Zimmerman, 1994, 1995, 1998). There has been a great deal of debate on how to handle outliers around what constitutes an outlier and whether it should be removed or not. Osborne (Osborne & A., 2004) proposed that "what to do" depends on "why an outlier is in the data" in the first place. Generally, when outliers are *illegitimately* included in the data, they should be removed (Barnett & Lewis, 1994); however, when the outlier is *legitimate*, it was suggested that they should be kept in, so as to obtain the most honest estimate of population-parameters possible (Barnett & Lewis, 1994).

In order to identify and understand the constitution of the outliers, the source consisting of outliers in subjective rating was visually inspected using Boxplot (see Figure 5.1) and it was found that the most extreme values were from the following subject:

- The subject number: 40; Age Group: #1; Male; Experience: 10 years; Annual mileage: less than 5000 miles.

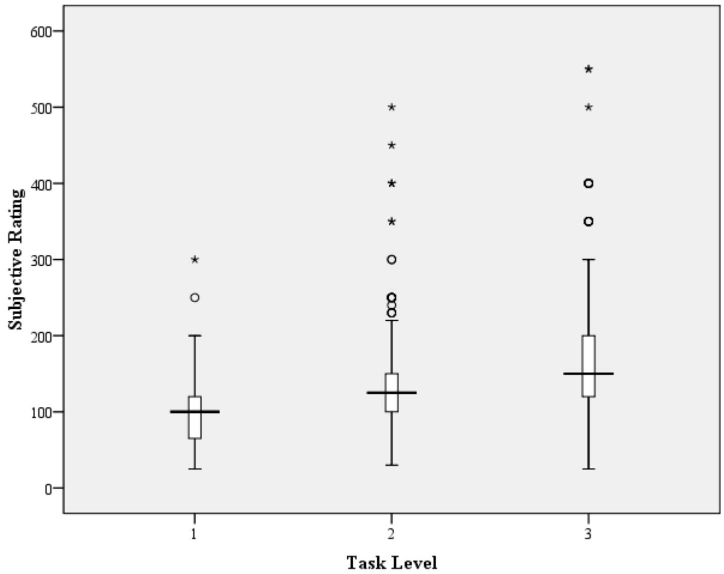


Figure 5.1 Boxplot for Subjective Rating across task difficult levels

(Note: a circle denotes outliers that are further than 1.5 x inter-quartile but within 3 x inter-quartile range, and a star denotes outliers that are farther than 3 x inter-quartile range.)

After having identified the major source of the outliers, it was clear that they were a legitimate part of the database, and therefore should be included. To deal with the violation against parameter analysis this caused, a non-linear transformation was conducted in order to normalise the distribution, as the data of SR were heavily skewed. A *Square Root Transformation*, recommended by Osborne (Osborne, 2002) has been conducted to adjust the SR distribution. The new variable was named as Trans_SR, which was calculated by the following formula:

$$\text{Trans\_SR} = \sqrt{\text{SR}}\ .$$

After transformation, the data became normal and were suitable for the parameter statistical method. However, like all transformed data, the process makes data hard to interpret directly. Therefore, the analysis of SR was still presented in the following sections, and the results from Trans_SR will only be reported if a different conclusion from that of SR could be drawn (e.g. one of the factors showed a significant effect on Trans_SR but not on SR, or vice verse).

The SR did not show a statistical difference between the task types, i.e., between auditory and visual tasks and the driving scenarios, i.e., between CF and FD ($p > 0.05$). The rating was therefore analysed as from the same population. The following effects on the subjective rating were investigated in this section: task difficulty (or task level), age, gender, driving experience, annual mileage (or exposure to driving), and the self-reported experience of, and the attitude towards, in-vehicle information systems (IVISs).

5.4.2 The Effect of Task Level on SR

The Levene's Statistical Test did not pass the homogeneity of variance assumption required by both ANOVA and non-parameter tests. The Non-parameter Test or Robust Test (e.g. Welch and BF Test) were therefore used to analyse the effect of task levels on SR. The results showed that the task difficulty level had a significant effect on the subjective rating (Welch Test, $F(2, 726) = 133.325$, $p<0.001$ and BF Test, $F(2, 865) = 129.186$, $p<0.001$). The Post Hoc Test with Bonferonni adjustment applied (Abdi, 2007) suggested that the differences existed between all three levels. The mean values of SR in each level were given by: Level 1 - 92.07 (SD: 42.789), Level 2 - 126.59 (SD: 61.657), and Level 3 - 170.40 (SD: 91.537), where the SD in brackets denotes the standard deviation of SR in each level.

The subjective rating increased steadily with the increasing task difficulty levels, see Figure 5.1. Also, a steadily increased standard deviation may be seen, which suggests a more remarkable individual difference while the task complexity increases.

5.4.3 The Effect Driver Characteristics on SR

This Section investigates the effects of age, gender, driving experience, and annual mileage on the subjective rating.

5.4.3.1 Gender

Generally, female drivers rated the workload by 9% higher than male drivers, with in average 136.88 for females and 125.92 for males (Welch Test and BF Test, $F(1, 964) = 6.796$, $p = 0.009$). The results are shown in Figure 5.2, in which "1" (with blue in colour) indicates male and "2" denotes female (with green in colour).

The subjective ratings of the workload from male and female drivers were then compared for each level of task difficulty. It was found that the major difference between the two genders on SR existed in level 1 ($p = 0.004$). In level 2, females still rated higher, but the effect was marginal ($p = 0.053$). In level 3, no significant difference was observed any more ($p = 0.192$).

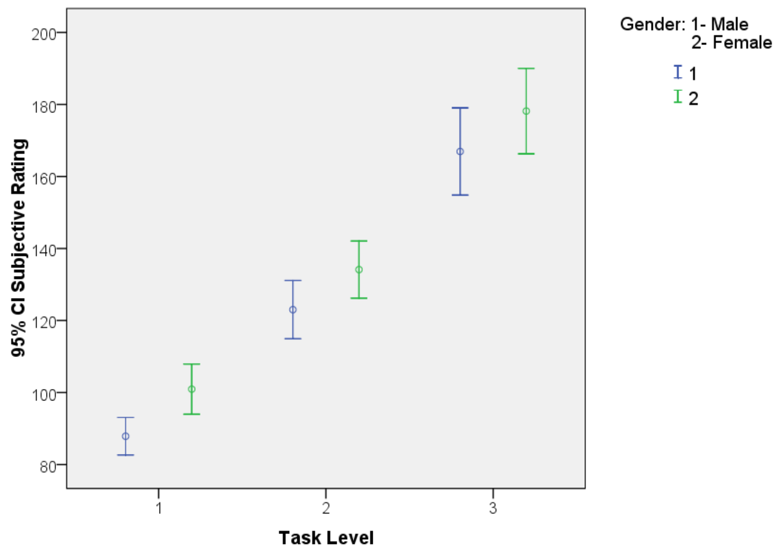


Figure 5.2 The subjective ratings for male and female drivers at different task levels

5.4.3.2 Age and Experience

Experience showed significant effect on SR, where the least experienced drivers (Group #1) showed a significantly lower subjective workload (Trans_SR, $F(2, 1184) = 3.787$, $p = 0.023$). This effect was also significant in each task level. Generally, Group #1 (123.54) was 6% lower than Group #2 (131.04) and 7% lower than Group #3 (132.15). The subjective rating compared between driving experience groups is shown in Table 5.6.

Table 5.6 The subjective rating with respect to the task levels and experience groups

Task Level	1			2			3		
Experience Group	# 1	# 2	# 3	# 1	# 2	# 3	# 1	# 2	# 3
Mean	**88.5**	90.49	97.65	**121.12**	128.1	129.23	**162.9**	174.1	171.25
SD	**44.39**	49.7	27.05	**59.271**	75.945	33.027	**85.541**	110.278	58.365

For effect of age on SR, as the equal-variance assumption was violated, the results from Robust Tests were reported. The results showed a difference in the subjective rating between the three age groups (Welch Test, $F(2, 756) = 10.751$, $p < 0.001$; BF Test, $F(2, 884) = 8.122$, $p < 0.001$). The Post Hoc Test suggested that the subjective rating in age group #1 was lower than that of age groups #2 and #3. In each level, the same effect still held true, i.e. the age effect existed in each level ($p < 0.05$), and also for all levels. The values of SR with respect to the task levels and age groups are shown in Table 5.7.

Table 5.7 The subjective rating with respect to the task levels and age groups

Task Level	1			2			3		
Age Group	# 1	# 2	# 3	# 1	# 2	# 3	# 1	# 2	# 3
Mean	**82.66**	93.98	100.17	**114.63**	135.50	130.27	**156.35**	179.63	175.65
SD	39.20	53.23	31.82	45.57	88.07	37.30	66.85	127.23	63.74

5.4.3.3 Annual Mileage

Similar effects were found for annual mileage groups, where the least experienced group reported the lowest workload level (59.47 versus 69.74 for Group #2 and 93.24 #3 averagely). The Robust Test showed a significant effect due to the annual mileage on the subjective rating (Welch Test $F(2, 690) = 24.609$, $p<0.001$; and BF Test $F(2, 831) = 27.437$, $p<0.001$). This effect was more noteworthy when tasks became more difficult, i.e., in levels 2 and 3.

5.4.3.4 Experience of, and Attitude towards, IVISs

Drivers' experience of, and attitude towards using auditory and manual systems were collected from the drivers' screening questionnaire (see Appendix A). All the group numbers of these four driver characteristics were given in such a way that the higher the Group Number is, the less experience of IVISs before, or more reserved attitude they have (i.e. considered IVISs as very distracting/unsafe). For example, for the experience of IVISs (2 groups in total), Group #1 includes subjects who claimed to have more experience of using such IVIS systems, and Group #2 contains subjects who less often or never use IVISs before. For the attitude Groups (3 groups in total), Group#1 consists drivers who thought the IVISs are relatively safe to use, while Group #2 represents a "neutral" attitude, and Group #3 are the drivers who were more concern about the distraction caused by them. A detailed description of how the groups were defined can be found in Chapter 3.

The Experience of using auditory systems had no effect on the subjective rating on auditory tasks ($p > 0.05$). The subjects who thought such systems are not distracting (Group #1) had a significantly higher subjective rating than the other two groups (see Figure 5.3).

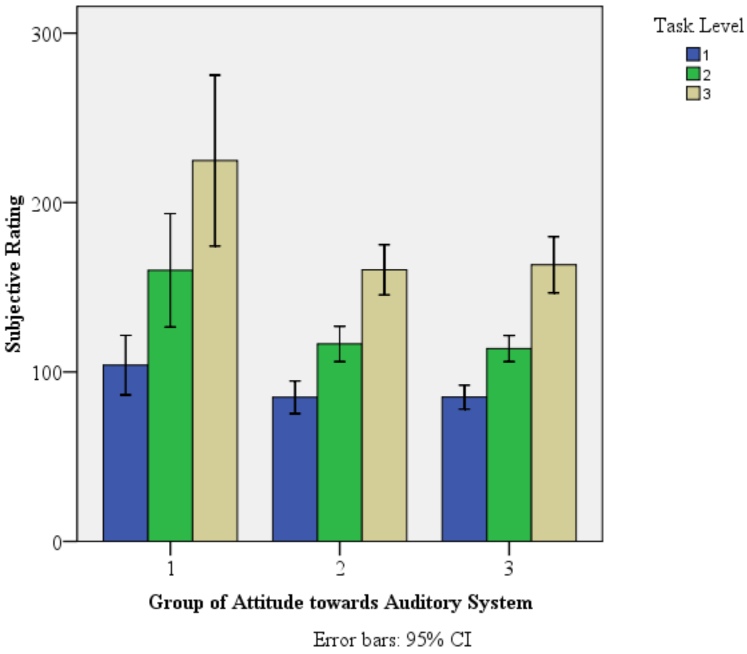


Figure 5.3 Subjective rating - by the group of attitude towards auditory systems

Similar story was found for experience of and attitude towards manual systems. The Experience of using manual systems had no effect on the subjective rating ($p > 0.05$). However, the group with subjects who thought such systems is safe (Group #1) showed a higher subjective rating (Atti_Visu Group #1 137.78 versus Group #2 116.79 versus Group #3 128.01).

5.4.4 Summary of Subjective Rating

There was no significant difference in subjective rating between road types and task types, i.e. drivers rated the two driving scenarios, and the visual and auditory tasks as equally demanding. The ratings of workload increases from task difficult level 1 to level 3, and all the increases across three task levels were significant ($p < 0.01$). Female drivers rated their experienced workload as higher than their male counterparts. This effect was more noteworthy in level 1 and slowly faded as the difficulty levels increased. In level 3, the difference of SR between the two genders was not significant any more.The drivers with less driving experience had a lower subjective workload. Similarly, the groups with younger age and least annual mileage showed a lower subjective rating on workload. Experience of using IVISs generally had no impact on drivers' subjective workload. The results from drivers' attitude towards using these systems whilst driving suggested the ones who thought using these are not distracting or safe rated their workload as higher the ones said they are unsafe.

5.5 Secondary Task Performance (STP)

5.5.1 Auditory Task Performance

As mentioned before, the secondary task performance (STP) for auditory tasks was measured by *"correct percentage"*, i.e. the percentage of correct counting(s) in 1, 2 or 3 audio target(s). In level 1, there was only one target sound, therefore the correct percentage would be 100% if the answer given by the subject was correct, and the correct percentage would be 100% if the answer was right; while in level 2, one correct and one wrong answer in two targets would be recorded as 50%.

As there was no statistical significant difference in the correct percentage between route types, the data collected from the Car-Following and Free-Driving Scenarios were combined into one dataset. Based on these data, the effects of task level, gender, age/experience, and the experience of, and attitude towards using in-vehicle auditory systems were investigated.

5.5.1.1 Task Level

A significant effect of task difficulty level was found for the auditory tasks performance from the reduced correct percentage (Welch Test $F (3, 386) = 33.616$, $p < 0.001$; and BF Test, $F (4, 567) = 34.965$, $p < 0.001$). The Post Hoc tests suggested that the correct percentage decreases significantly for all 3 levels (Tukey HSD Test, $p < 0.05$). Figure 5.4 shows the comparison of the correct percentage between task levels, where the error bar represents the 95% Confidence Interval (CI).

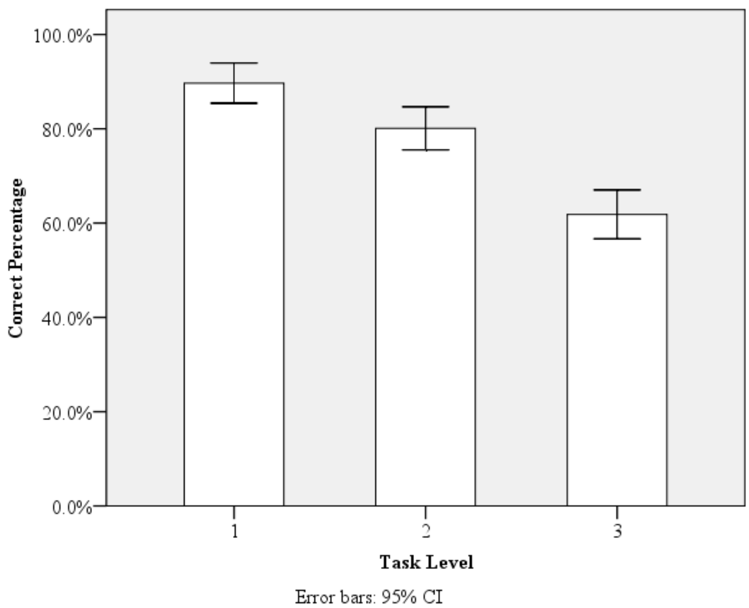


Figure 5.4 The correct percentage of the auditory secondary tasks across task levels

5.5.1.2 Gender

Averagely, female drivers performed the auditory tasks in a lower correct rate by 8% than male drivers (Female 73.17% versus 79.39% for male, ANOVA $F(1, 586) = 4.037$, $p = 0.045$). The effect became more significant when the task difficulty increased. In level 1, there was no significant difference between the two genders, $F(1, 195) = 0.644$, $p = 0.423$; while in level 2, the difference became noticeable, $F(1, 199) = 4.475$, $p = 0.036$; and the gender difference was even larger in level 3, $F(1, 188) = 5.943$, $p = 0.016$. The comparison is shown in Figure 5.5.

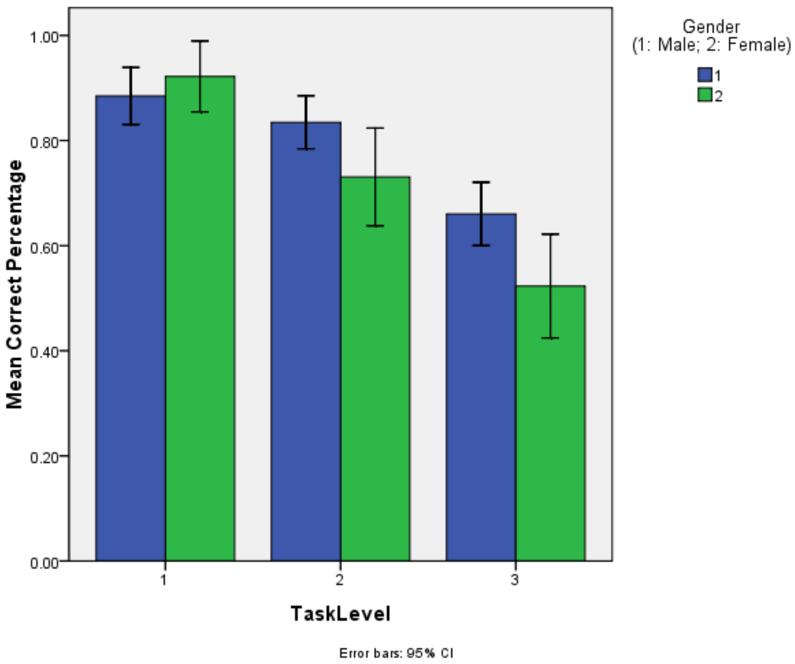


Figure 5.5 The correct percentage of the auditory secondary tasks, by gender

5.5.1.3 Experience/ Age and Mileage

Experience had moderate effect on the correct rate for auditory tasks (Chi2 (2) = 4.763, $p = 0.092$). It may be observed from Figure 5.6 that Experience Group #1 (less experienced group) had a lower correct rate than Group #2 and #3 (75.55% versus 83.38% for Group #2, and 81.72% for Group #3, on average), i.e. 8% lower than Group #2, and 6% than Group #3. Experience Group #2 had lower correct rate for auditory tasks comparing to the other two groups. But this difference was not statistically significant ($p > 0.05$). Age showed similar effect, with Group #1 had lower correct rate than Group #2 and #3. Annual mileage did not impact the correct rate of performing auditory tasks.

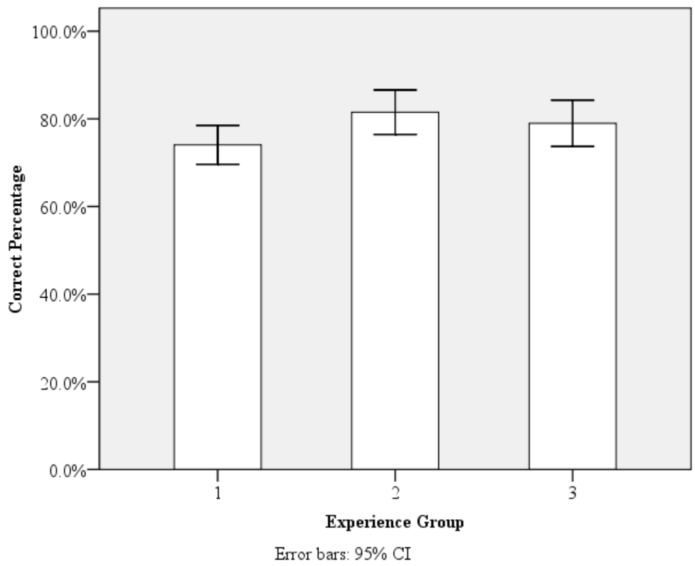


Figure 5.6 The correct percentage of the auditory secondary tasks, by experience

5.5.1.4 Experience of, and Attitude towards using Auditory Systems

Experience of and attitude towards being auditory distracted whilst driving had no impact on the correct rate of performing auditory tasks. On average, the correct rate for less experience group (Expe_Audi Group #1) was 77.2%, and 77.55% for more experienced one (Group #2). But from Figure 5.7, a lightly lower correct rate from the group which thought mobile phone was more distracting (Atti_Audi Group #3) can be observed, especially for the task level 3. On average, the correct rate for drivers who thought them as not distracting (Atti_Audi Group #1) was 79.09%, 80.87% for Group #2, and 73.46% for Group #3.

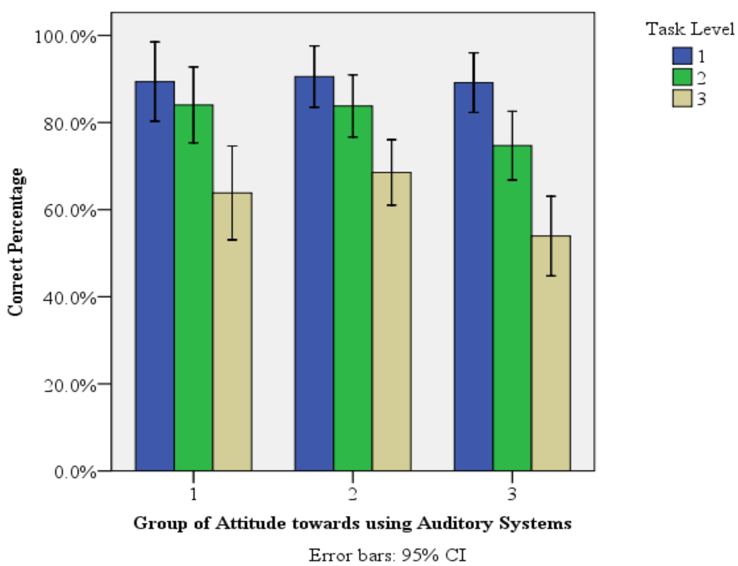


Figure 5.7 The correct percentage by groups of attitude towards using auditory tasks systems

5.5.2 Visual Task Performance

The visual task performance was assessed by several parameters; the Reaction Time (RT) to each image, which is the average of the time lags from the appearance of each new image to the first correct click on the touch screen (i.e. click on the "target"); the Standard Deviation of Reaction Time (SDRT), which is the deviation of RT for each individual visual task; the Error Percentage (ErP) which describes the percentage of mistaken clicks in each task; and the Error Distance (ErD) which is the mean distance from the nearest "target circle" when a missing-target click occurs. All the visual task performance data were explored first, in order to decide later which statistical test should be used. The data of RT and ErD were heavily skewed, because there is a natural under limit - for example, Reaction Time cannot lower than a certain value for human beings, which is why the RT data distribution is skewed to the right. Figure 5.8 shows the data distribution of RT when performing the visual tasks in Car-Following, where the red line shows the fitness curve for a lognormal distribution ($p > 0.05$).

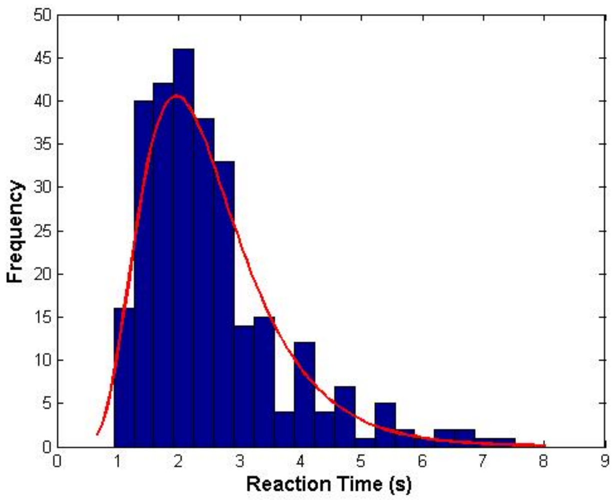


Figure 5.8 The distribution of Reaction Time when performing the visual tasks in CF

The RT and SDRT had equal variances across two scenarios, while unequal variances were found for ErP and ErD. Taking the distribution of each parameter into account, for RT and SDRT, the results from Mann-Whitney U Test were reported; while for ErP, ErD and Trans_ErD, the Welch Tests were conducted. The results showed that RT, SDRT and ErP measures differed significantly between the two test scenarios, with longer RT and higher ErP in the Free-Driving Scenario. Figure 5.9 shows RT in the two scenarios. The analysis was then conducted separately in two conditions: Car-Following (CF) and Free-Driving (FD).

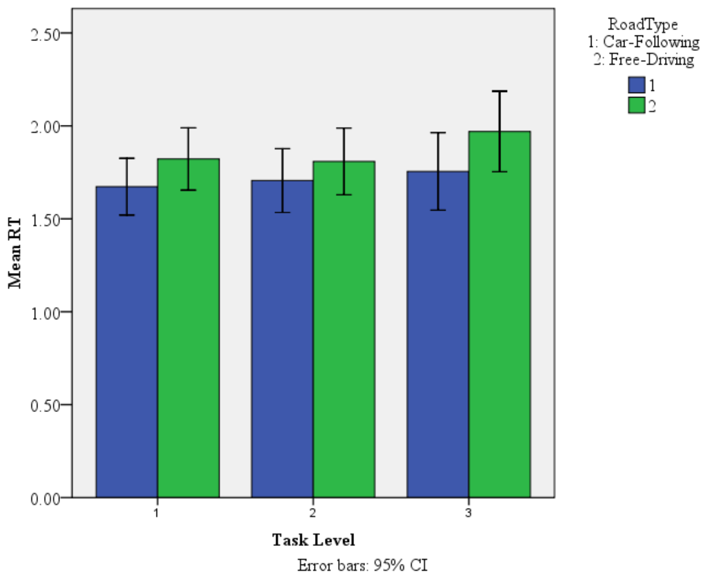


Figure 5.9 Reaction Time at different task levels – compared between Scenarios

5.5.2.1 Secondary Task Performance in Car-Following

5.5.2.1.1 Task Level

In Car-Following, the reaction time increased with the task difficulty (Chi^2 (2) = 13.554, p = 0.001. The Post Hoc analysis suggested that drivers had a longer reaction time in task level 3 (M = 2.28, 95% CI [2.08, 2.47]) than that in task level 1 (M = 2.82, 95%CI [2.57, 3.07]). It was also observed that SDRT in level 1 was significantly higher than that in levels 2 and 3. No statistically significant difference across the task levels on ErP and ErD was found (p>0.05), similar with the findings in HASTE (Engström, Johansson & Östlund, 2005). The RT, SDRT, ErP and ErD across all three levels are illustrated in Table 5.8.

Table 5.8 Visual task performance in Car-Following compared between levels

Task Level		RT (s)	SDRT	ErP	ErD
1	Mean	2.28	1.08	12.91%	8.83
	SD	0.99	0.70	12.06%	12.33
2	Mean	2.45	0.74	11.82%	9.86
	SD	1.18	0.63	9.63%	13.08
3	Mean	2.82	0.91	11.64%	8.74
	SD	1.29	0.82	11.32%	12.95

5.5.2.1.2 The Effects of Gender, Age, Experience and Mileage

5.5.2.1.2.1 Gender

On average, compared to their female counterparts, the male drivers reacted faster to the visual secondary tasks by 23.29% (RT for male: 2.34s; RT for female: 2.89s), with a 29.81% higher percentage of making mistakes (ErP for male: 13.42%; ErP for female: 9.42%) in

Car-Following ($p < 0.05$). This effect mainly existed in ErP in level 1, and in RT and ErP in level 3 ($p < 0.05$), although this was not significant. No gender effect was found for either SDRT or ErD. Table 5.9 demonstrates the RT and ErP compared between male and female drivers at different task levels.

Table 5.9 Visual task performance in Car-Following compared between gender

Task Level	Gender	RT	SDRT	ErP	ErD
1	Male	2.21	1.09	14.84%	9.46
	Female	2.41	1.04	8.86%	7.53
	F	0.691	0.106	6.875	0.542
	df	1, 45.996	1, 100	1, 79.763	1, 100
	p-value	0.410	0.745	0.010**	0.463
2	Male	2.26	0.69	12.23%	10.87
	Female	2.85	0.85	10.98%	7.78
	F	3.792	1.244	0.376	1.241
	df	1, 39.375	1, 49.714	1, 99	1, 99
	p-value	0.59	0.270	0.541	0.268
3	Male	2.55	0.84	13.18%	9.04
	Female	3.39	1.05	8.42%	8.13
	F	6.944	1.526	5.731	0.109
	df	1, 40.979	1, 100	1, 95.248	1, 100
	p-value	0.012**	0.220	0.019**	0.742
Total	Male	2.34	0.87	13.42%	9.78
	Female	2.89	0.98	9.42%	7.81
	F	10.114	1.380	1, 250.534	1.594
	df	1, 126.280	1, 175.082	11.034	1, 303
	p-value	0.002	0.261	0.001	0.869

Note: 1. the Diff in the table was calculated by:
(value from female – value from male) / value from male.
2. ** denotes a significant effect, where p<0.05.

5.5.2.1.2.2 Driving Experience, Age and Annual Mileage

Results showed that the driver's age and experience were sensitive to almost all parameters of the visual task. The general trends can be described as following: the most experienced driving Group (#3) had the shortest reaction time, but made more mistakes (SNK and Tukey HSD, $p < 0.05$). This could be due to their higher skill and a more relaxed attitude, as suggested by their Post-Driving Interviews, in which they said they prioritised driving over the secondary tasks, and only performed them when they could. Experience Group #2 showed the best secondary task performance, with shorter reaction time, fewer mistakes and less error distance. The drivers with less than 10 years' experience had slow reaction to the visual tasks (10.36%

slower than Experience Group #3), and they also made mistakes to a larger extent (22% larger than Experience group #3), see Table 5.10.

Table 5.10 Secondary task performance for different experience groups in Car-Following

Experience Group	RT	SDRT	ErP	ErD
# 1	**2.66**	1.06	0.12	**10.88**
# 2	2.38	0.67	**0.09**	**6.31**
# 3	2.43	0.92	0.15	9.40

Age Group #3 had a much shorter RT than Age Group #2, a higher ErP and higher ErD. Age Group #2 showed a much longer RT, but lower ErP and ErD. Age Group #1 had a medium RT, lower ErP and higher ErD. The impact of Annual Mileage showed a similar trend, with the Group having the least mileage showing longer RT but lower ErP, while the Group with most mileage showed the fastest reaction, but higher ErP (SNK and Tukey HSD, $p < 0.05$).

5.5.2.1.3 Experience of, and Attitude towards, Using Manual IVISs

The group of drivers who had more experience in using manual IVISs had a lower ErP (ErP for more experienced IVIS group: 9.36%; ErP for less experience IVISs group: 12.75%). This suggests that previous experience of using manual IVISs improves the accuracy ($F(1, 303) = 7.273$, $p = 0.007$). This group also showed a lower SDRT ($F(1, 303) = 4.790$, $p = 0.029$). Driver's attitude towards manually using IVISs while driving showed that drivers who were more against using manual IVISs (Group #3) had short RT to the secondary tasks ($F(2, 302) = 3.539$, $p = 0.030$), with more mistakes ($F(2, 302) = 6.879$, $p = 0.001$) than the other two groups. The effects of experience and attitude on manual IVISs in Car-Following on visual task performance are listed in Table 5.11.

Table 5.11 STP by groups of experience and attitude on manual IVISs in Car-Following

Expe_Visu	ErP	SDRT	Atti_Visu	RT	SDRT	ErP
Group #1	9.36%	0.87	Group #1	2.64	0.88	11.08%
Group #2	12.75%	1.05	Group #2	2.81	1.02	8.39%
			Group #3	2.33	1.04	14.65%

5.5.2.2 Visual Task Performance in Free-Driving

5.5.2.2.1 Task Level

In the Free-Driving Scenario, the task level has significant effects on RT ($F(2, 294) = 4.781$, $p = 0.009$) and SDRT ($Chi^2(2) = 8.587$, $p = 0.014$), but has no effect on ErD ($F(2, 294) = 2.011$, $p = 0.136$) or ErP ($F(3, 294) = 0.648$, $p = 0.524$). The Post Hoc analysis revealed that RT in task level 3 was longer than that in levels 1 and 2, and SDRT in level 1 was higher than that in levels 2 and 3 (Tukey HSD Test, $p < 0.05$). The value of visual task performance for each level is shown in Table 5.12.

Table 5.12 Visual task performance for each level in Free-Driving

Task Level		RT(s) *	SDRT*	ErP	ErD
1	Mean	2.64	1.25	17.17%	15.96
	SD	1.07	0.79	14.45%	23.02
2	Mean	2.71	1.01	15.05%	12.82
	SD	1.06	0.74	12.30%	16.52
3	Mean	3.10	1.00	15.60%	10.18
	SD	1.20	0.74	13.85%	14.05

Note: * denotes that a significant effect exists, where $p < 0.05$.

5.5.2.2.2 The Effects of Gender, Age/Experience and Annual Mileage on STP in Free-Driving

5.5.2.2.2.1 Gender Effect of STP in Free-Driving

Both generally and in each task level, there was no significant gender effect on any of the measurements in the Free-Driving Scenario, but the results showed a similar trend with that in the Car-Following, i.e., female drivers had a longer RT, lower ErP, and less ErD.

Table 5.13 Visual task performance in Free-Driving compared between gender

Gender		RT(s)	SDRT	ErP	ErD
Male	Mean	2.79	1.11	16.59%	14.27
	SD	0.99	0.80	14.06%	20.54
Female	Mean	2.87	1.03	14.52%	10.16
	SD	1.39	0.69	12.29%	11.86
Diff		2.87%	-7.21%	-12.48%	-28.80%

5.5.2.2.2.2 Experience/Age Effect in Free-Driving

In Free-Driving, the effect of driving experience STP showed similar findings with that in Car-Following. RT steadily decreased with the increased experience, especially for the experience group #1, who had a much longer RT ($F(2, 292) = 3.705$, $p = 0.026$). Again higher ErP ($F(2, 292) = 21.682$, $p < 0.001$) was also found for the most experience group (Group #3). Experience group #2 demonstrated a relatively short RT, the lowest ErP to the lowest extent (ErD), still showing the best STP in all three groups. For details, refer to Table 5.14.

Again, the same as that in Car-Following, the age group # 1 had the shorter RT ($F(2, 292)=10.153$, $p < 0.001$) and a lower ErP ($F(2, 292) = 7.483$, $p = 0.001$). The age group #2 showed a longer RT and a lower ErP. Age Group #3 reacted the quickest to the secondary task, but with higher ErP. No difference was found for the age effects on SDRT and ErD (Kruskal Wallis Test p>0.05), see Table 5.14.

Table 5.14 RT, SDRT, ErP and ErD with different age and experience groups in Free-Driving

Experience	RT(s)	SDRT	ErP(%)	ErD	Age	RT(s)	SDRT	ErP(%)	ErD
#1	2.92	1.19	17.39	14.89	#1	2.66	0.99	14.28	11.30
#2	2.83	0.94	10.96	9.59	#2	3.17	1.17	14.54	13.20
#3	2.62	1.08	18.98	13.53	#3	2.61	1.09	19.02	14.43

In Free-Driving, Mileage Group #3 showed a shorter RT ($F(2, 292) = 7.130$, $p = 0.001$) and higher ErP ($F(2, 293) = 4.523$, $p = 0.012$). This is consistent with that found in Car-Following.

5.5.2.2.3 Effect of Experience and Attitude towards Manual IVISs on STP in Free-Driving

Same with that in Car-Following, it was found that previous experience of manually adjusting IVISs improves the task accuracy (ANOVA Test, $F(1, 294) = 14.719$, $p < 0.001$). The group of drivers who had more experience in using manual IVISs had a lower ErP (ErP for more experienced IVIS group: 12.07%; ErP for less experience IVISs group: 18.10%). Driver's attitude towards manually using IVISs while driving showed that drivers who were more against using manual IVISs had short RT to the secondary tasks (Trans_RT $F(2, 292) = 3.858$, $p = 0.022$) than the other two groups (RT for Group #1: 2.96s; Group #2: 2.99s; Group #3: 2.62s). The performance of visual tasks by different groups of experience and attitude on manual IVISs in Free-Driving is illustrated in Table 5.15.

Table 5.15 STP by different groups of experience and attitude on manual IVISs in Free-Driving

Expe_Visu	RT	SDRT	ErP(%)	ErD	Atti_Visu	RT	SDRT	ErP(%)	ErD
Group # 1	2.86	1.05	**12.07**	13.69	Group # 1	2.96	1.05	16.57	13.87
Group # 2	2.92	1.21	18 .10	17.11	Group # 2	3.00	1.13	12.24	17.00
					Group # 3	**2.62**	1.27	16.31	16.20

5.6 Summary and Discussion

The main findings from this Chapter are:

- The auditory and visual tasks designed for this study caused similar levels of workload on drivers, which provided a basis for comparing the effects of these different types of workload.
- According to drivers' own Subject Ratings, their workload increased with the task difficult level, but the workload between the two Driving Scenarios was comparable.
- In each Scenario, the performance of both auditory and visual tasks decreased as the task difficulty level increased.

- According to their Subjective Ratings, female drivers generally experienced a higher workload, and had a lower performance in the secondary tasks. However, they made fewer errors in the visual tasks, but with longer reaction times.
- Younger drivers and those with less driving experience rated their workload as lower, but showed significantly worse performance in both the auditory and visual tasks.
- Previous experience of using IVISs did not reduce drivers' workload, but improved their accuracy in performing the secondary tasks.
- Subjects who thought using IVISs were safe rated their workload as being higher. However, the impact of attitude towards IVISs was not conclusive, as those who thought IVISs were distracting or unsafe had a lower auditory task performance, but showed a faster reaction time for the visual tasks.

Details of these findings are discussed and explained below.

5.6.1 Effects of Secondary Task

For both the auditory and visual tasks, the level of difficulty had a pronounced effect on drivers' Subjective workload Rating and nearly all their secondary task performance measurements. As the level of task difficulty increases, the workload Rating increased accordingly, while secondary task performance decreased remarkably - for example, for the auditory tasks, the percentage of correct answers decreased by 31% from level 1 to level 3 overall, while for the visual tasks, the reaction time increased by 7% from level 1 to 2, and 15% between levels 2 and 3 in Car-Following, and by 3% from level 1 to 2, and 14% between levels 2 and 3 in Free-Driving. The ability of Subjective Rating to distinguish between task difficulty level was also found in previous research (Angell et al., 2006; Burns et al., 2002; Engström, Johansson & Östlund, 2005; Östlund et al., 2004), with a lower correct percentage (for auditory tasks) and longer reaction time (for visual ones) in the higher difficult levels (Engström, Johansson & Östlund, 2005; Östlund et al., 2004).

The deterioration in secondary task performance with difficulty level reflects not only the extra workload - and therefore effort - required in performing the more difficult tasks, but also that drivers may be *unwilling* or *unable* to devote further resources to compensate for their performance in higher difficulty tasks under real-road driving conditions.

5.6.2 Effects of Driver Characteristics

5.6.2.1 Gender Effects

In general, female drivers rated the workload of all tasks as being significantly higher than their male counterparts. They also showed a lower percentage of correct answers in performing the auditory tasks (73%) compared to males (79%), and a longer reaction time in visual tasks (2.89s versus 2.34s in Car-Following and 2.87 versus 2.79 in Free-Driving), although this was compensated by greater accuracy, with clicking error percentage of 9% versus 13% in Car-Following, and 15% versus 17% in Free-Driving.

For both auditory and visual tasks, female drivers had a slightly higher (or same) secondary task performance as males when the difficulty level was low, but their task performance dropped remarkably in the highest level, suggesting that the gender difference is more pronounced under condition of higher workload or “worse” driving condition.

5.6.2.2 Effect of Driving Experience/Age and Mileage

Driving experience also had a major effect on workload and secondary task performance. Drivers with the least experience (Driving Experience Group #1) rated their workload as lower than other Groups, yet they showed the worst secondary task performance, with the lowest percentage of correct answers when performing the auditory tasks (75.55% versus 83.38% and 81.72% for Group #2 and #3 respectively), and longer reaction times and larger clicking error distances for the visual tasks (for example, 2.66s and 10.88 pixels compared to 2.38s and 6.31 pixels for the best Group in Car-Following). Their poorer secondary task performance and lower workload Rating (6% lower than Group #2, and 7% lower than Group #3) suggest these drivers were less aware of the additional demand caused by in-vehicle tasks, and either possessed a higher risk tolerance to them, or had a *lower capacity* than the other Groups to cope with dual-tasking (see Chapter 6). It is also possible they were more concerned with the primary task of driving, and therefore de-prioritise the secondary tasks, as suggested by Wickens (Wickens, 1984a), or they were not significantly engaged in performing them.

Compared to the least experienced, the most Experienced Group (#3) showed a higher percentage of correct answers when performing the auditory tasks (6% higher than Group #1), and shorter reaction times in the visual tasks (e.g. 2.43s in Car-Following). A *slightly* better secondary task performance was found in Experience Group #2 (compared to #3), but this difference was not significant except for a lower error percentage and clicking distances when performing visual tasks. However, given drivers have limited resources, this performance difference could be a trade-off between different tasks, and therefore needs to be discussed in the context of their primary driving performance (see Chapter 6) and attention allocation or eye movement (see Chapter 7).

Drivers who had the lowest annual mileage rated their workload as lower than the other two groups, while those in the highest Mileage Group had the shortest reaction time. Higher mileage i.e. greater exposure to driving has been already been associated with driving performance improvements, and this evidence showed that experience in one task (i.e. primary driving) can help improve secondary task performance, as increased *automation* of the primary task can reduce the overall workload on drivers when dual-tasking.

Compared with driving experience and mileage, the effect of age is less significant in this study. Age Group #2 (31-40 year olds) rated their workload as lower than Group #1 (21-30 year olds) and #3 (41-60 year olds). This Group also showed a slightly higher auditory task performance and smaller visual clicking error distances than the other Groups, although the effect was not significant. The eldest Age Group (#3) showed shorter reaction time, but more errors in their visual tasks. Overall, the effect of age on secondary task performance showed few distinct trends compared to driving experience and mileage. As comparison (see Chapter 6), the effect of age on *primary* driving performance is more pronounced, but still less significant than experience and mileage.

5.6.2.3 Effect of Experience/Attitude towards Using IVISs

Previous experience in using IVISs had little effect on workload Rating and auditory task performing, but improved the accuracy of the visual tasks (3% less error than the Group with less experience). This suggests that the ability of accurately manipulating IVISs can be improved by practice, and more evidence was found in the experiment, where the clicking error distances reduced from the first time visual tasks was performed in Free-Driving to the third time.

Drivers who rated auditory IVISs systems as distracting (Atti_Audi Group #3) showed a lower performance in auditory task performance (6% lower correct rate than Group #1), while those who rated visual IVISs as unsafe (Atti_Visual Group #3) had a slightly shorter reaction time (e.g. 2.33s versus 2.64s in Free-Driving), with higher clicking error percentage (14.65% comparing to 11.08% for Group #1 in Car-Following). Interestingly, comparing the Subjective workload Rating between drivers' attitudes towards using IVISs showed contradictory results to expected, i.e. those who thought using IVISs as safe rated the workload as being *higher* during the experiment. More evidences is therefore needed to explain how drivers approached the primary and secondary tasks between the different groups (see Chapter 6), and how they allocated their attention and eye movement when dual-taking (Chapter 7).

5.6.3 Comparison between Scenarios

Overall, during the experiment, there was no significant difference in the Subjective workload Rating and secondary task performance between Car-Following and Free-Driving, although the two Scenarios are remarkably different, i.e. keeping a constant headway in busy traffic conditions in one, versus keeping a constant speed while driving through quiet road sections in the other. Some differences were found in drivers' visual task performance, with a longer reaction time, higher error percentage and clicking error distance observed in the Free-Driving Scenario. Given drivers did not report a higher workload for this Scenario, the difference to Car-Following could be explained by the perceived need for higher *and* constant speeds involved in Free-Driving. It has been suggested (Senders et al., 1967b) that drivers tended to drive at a speed which balances their information processing capacity with the rate at which such information is generated from the road. When instructed to drive at the higher and constant speed in Free-Driving, it is possible their visual capacity is more focussed to deal with the correspondingly higher rate of road information generation than on the need to process other (secondary) tasks. Therefore, with fixed information processing capacity, when more attention is required for the primary task of driving, secondary tasks are de-prioritised, simplified or ignored.

The workload Rating and secondary task performance between the two Scenarios were broadly similar in terms of the effects of different driver characteristics, although some secondary task performance parameters were more sensitive to the Car-Following Scenario than in Free-Driving - for example, the effect of gender was observed in both Scenarios (e.g. with a longer reaction time), but was more pronounced in Car-Following. Therefore the key determinants of secondary task performance are these characteristics, which are not dependent on the driving Scenario and circumstances.

Chapter 6 Impact of Secondary Tasks on Driving Performance

6.1 Introduction

This Chapter presents the *primary* performance of drivers when conducting the secondary tasks, compared to baseline driving. The effects of different driver characteristics including age, gender, driving experience, annual mileage, experience and attitude towards using IVIS on primary driving, while performing the secondary tasks were also investigated and reported in this Chapter. The impact of leading and/or tailing vehicles on driving performance was eliminated from the analysis.

In a driver-vehicle system, the primary and secondary task demands make up the input, while the vehicle dynamic is the output (see Chapter 2). The secondary tasks may result in both intended and unintended changes in driving behaviour, which may further affect the driving efficiency and safety. Based on the hierarchical model of driving behaviour proposed by Michon (Michon, 1985), the workload increased due to the secondary task is most likely to affect the driving behaviours at the manoeuvring and control levels, which correspond to the performance of longitudinal controls, and lateral controls.

This Chapter presents the primary driving performance when drivers were dealing with the pre-designed auditory and visual tasks in both test driving scenarios – Car-Following and Free-Driving. In the same way as introduced in Chapter 5, the analyses have been conducted in the following four *task and scenario* types:

5. Auditory Task when Car-Following (ATCF);
6. Visual Task when Car-Following (VTCF);
7. Auditory Task when Free-Driving (ATFD);
8. Visual Task when Free-Driving (VTFD).

The comparisons of driving performance were made between four levels: three levels of task difficulty and baseline driving (i.e. normal driving without performing any extra task), along with the effect of driver characteristics, such as gender, age/experience, drivers' experience of and attitude towards using IVISs, the analyses will lead to the conclusion of to what extent performing in-vehicle secondary tasks can impact on driving performance, and how the factors mentioned above may influence any deterioration.

6.2 Measures of Driving Performance

The performances of individual drivers are evaluated using a number of longitudinal and lateral control parameters taken from the vehicle, which reflect manoeuvring and control of driving behaviour. These parameters or driving performance measurements are illustrated in Table 6.1.

Table 6.1 Measurements of driving performance

Measurements on the Longitudinal Controls	
Free-Driving	Mean speed
	Standard deviation of speed
	Maximum speed
	Mean acceleration/deceleration
	Maximum acceleration/deceleration
Car-Following	Mean distance headway
	Standard deviation of distance headway
	Minimum distance headway
	Mean time headway
	Standard deviation of time headway
	Minimum time headway
	Time-To-Collision (TTC)
Measurements on the Lateral controls	
Steering	Standard deviation of steering wheel angle
	Steering wheel Reversal Rate (RR)
Lane Keeping	Mean lateral position (MSLP)
	Standard deviation of lateral position (SDLP)

From Table 6.1, the *longitudinal control* measures can be classified into two types according to the driving Scenarios of Free-Driving and Car-Following. In Free-Driving, *speed* plays an important role in the risk of accidents: the higher the speed, the increased risk of accidents, and a power function has been suggested by (Aarts & van Schagen, 2006b) to describe the relationship between speed and accident risk. As described in the same paper, the speed *deviation* is also closely related to accidents because the risk increases with more variations, and hence the deviation in speed is also considered as an indicator for road safety. In the Car-Following scenario, the distance/time headway provides a better measure of the safety margins than speed, and (Vogel, 2003) for example, demonstrated that small time headways result in higher potential danger. According to Summala's results (Summala, 1981) in an actual driving situation, drivers' average reaction time is around 2.5 seconds, yet 2 seconds is the recommended minimum time headway in most driving manuals. Even worse, most drivers tend to keep a headway shorter than the recommended minimum, and short headways are therefore fairly common, but still considered to be much riskier (Evans, 1991).

As shown in Table 6.1, the *lateral* control parameters can be classified into two groups by steering and lane-keeping behaviour, in which the measures of steering represent the rate of steering wheel adjustments and the measures of lane-keeping indicate the increased accident probability. Previous research (Johansson et al., 2004) showed that, when the steering wheel adjustments increase, especially small steering corrections, the effort spent on the lateral control task (e.g. on a narrow road) will also increase. The standard deviation of lateral position (SDLP) is on one of the most consistent and commonly used measures to evaluate the lane keeping and in-vehicle technology, which historically used as a well-established proxy for risk, i.e., the probability of the vehicle getting out of the lane (Janssen et al., 2004; Wewerinke & Hogema, 2003). It was expected that SDLP would increase when performing In-Vehicle secondary tasks, because of the *decreased* attention to the primary driving task.

6.2.1 Measurements for Longitudinal Control

The following measures of longitudinal controls were used as measures for driving performance and further analysis:

- **Mean speed** (MNSP) and **speed deviation** (SDSP): In the Car-Following scenario, the speed deviation was the standard deviation of speed. While in the Free-Driving Scenario, as the instruction given to drivers was to maintain a constant speed at 55 miles per hour (i.e. 88 km/h), therefore the speed deviation was calculated using the following formula instead:

$$SDSP = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_i - 88)^2} \qquad (6.1)$$

where n is the number of the readings in each case.

The deviation in the Free-Driving Scenario still used the abbreviation of SDSP for convenience.

- **Mean** and **deviation** of **distance headway** (MNHW and SDHW), mean and deviation of **time headway** (MNTHW and SDTHW), the **minimum** distance and time headways (MinHW and MinTHW). These parameters are important in the Car-Following Scenario, because drivers were instructed to maintain a constant distance from the vehicle in front, and therefore any major change or deviation can be an indicator of driving performance deterioration. These headway parameters are not relevant to Free-Driving.

6.2.2 Measurements for Lateral Control

In terms of lateral controls, the following measures were computed and analysed:

Standard Deviation of Lane Position (SDLP): SDLP presents the extent to which drivers are able to keep a stable position in the lane they are driving in, i.e. the variation in lane position; therefore a lower SDLP value denotes a more stable lane control. The lane position was calculated from the left and right lane offsets, which are the distances from the centre of the IV to the left and right lane markings respectively. The IV records left lane offsets as negative values and right as positive, lane position was therefore calculated by:

$$\text{Lane Position} = \frac{\text{Right Lane Offset} - \text{Left Lane Offset}}{2} + \text{Left Lane Offset} \tag{6.2}$$

Figure 6.1 shows an example the SDLPs for the three task types (baseline, performing auditory and visual task) in the Free-Driving Scenario for subject #5. It can be seen from the figure that that the lane keeping was the most stable in baseline driving, and less stable when performing the visual task, whereas the SDLP value was nearly twice as much as for the baseline.

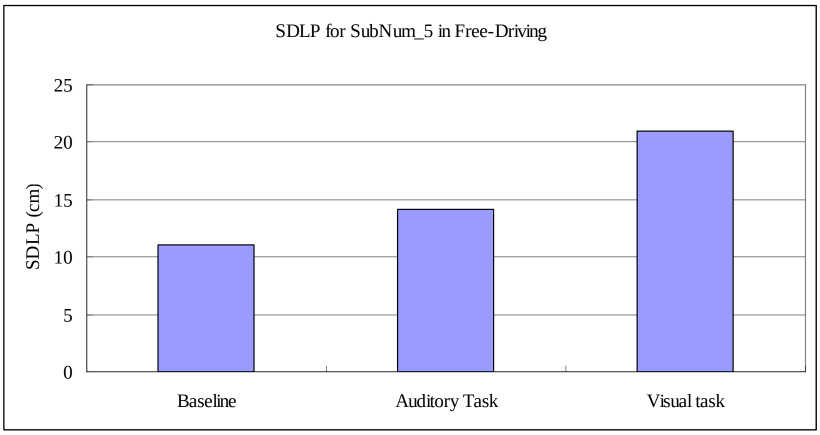


Figure 6.1 SDLPs of three task types for Subject #5 in Free-Driving

Steering Behaviour was assessed by the steering wheel **Reversal Rates** larger than 1.5 and 3, 5, 6, 7, and 8 degrees (i.e. the number of steering wheel adjustments per minute which are larger than these degree) or RR1.5, RR3, RR5, RR6, RR7, RR8; and **Steering Entropy** (SE).

A pre-analysis was conduced to decide which Reversal Rate parameters were best to describe driving performance. The analysis results showed that the reversal rate larger than 3° (RR3) was very similar to that of RR4, 5, 6, 7, and 8; and showed a better homogeneity in each task type. Hence in this thesis, the results of RR4, 5, 6, 7, and 8 will not be reported, given that they do not contradict the findings from **RR3** and **RR1.5**, namely *large* and *minor* steering wheel adjustments. However, by doing this, it is possible that significantly large fluctuations in steering adjustments, e.g. RR7 or RR8, which can be a sign of driving performance deterioration, will sometimes be dampened or smoothed. But given SDLP is a more consistent measure of performance deterioration, the benefit of analysing RR larger than 3 degrees separately was deemed superfluous. It should be noted that when workload increases, drivers are often less able to control their vehicle, which can result in increased steering wheel adjustments. However, these adjustments are significantly large and tend to be abrupt in nature, i.e. larger than three degrees, and peak patterned. When the adjustments are relatively smaller (including RR1.5 and RR3), the increases accompanied by comparable SDLP reflect the extra effort of drivers to adjust their steering to keep lane position.

In addition, Steering Entropy can be another indicator for driving performance deterioration, as it has been well proved (Boer, 2000; Nakayama et al., 1999) to be sensitive to steering behaviour changes, indicating unpredictability of steering, i.e. the *higher* the value, the less predictable the steering. Based on the definition by Nakayama et al. (Nakayama et al., 1999), Steering Entropy is related to the steering *angle*. In this research, the steering angles a_n are recorded at a frequency rate of 10 Hz, which is larger than the lowest sampling frequency that could be used to reveal a human operator's control response to manual tracking tasks (Nakayama et al. 1999). Based on the measurement data of the steering angles a_n, a second-order Taylor expansion on time is used to predict the steering angle at any given time. After a simple derivation, the steering angle b_n can be predicted using three subsequent points in the measurement data:

$$b_n = \frac{5}{2} a_{n-1} - 2a_{n-2} + \frac{1}{2} a_{n-3},$$

and the prediction error is calculated by:

$$e_n = b_n - a_n$$

The prediction error e_n can be assumed to be normally distributed. Therefore the mean prediction error is obtained using all error data within 1.2 standard deviations, to encompass 90% of the population. Suppose 90 percent of the population falls between -*C* and *C*, then the prediction error distribution between -*C* and *C* can be divided into nine bins, from which the steering entropy is derived (Nakayama et al. 1999):

$$H_p = -\sum_{m=1}^{9} p_m \log_9 p_m$$

in which p_m is the probability of being in the *m*th bin. Steering Entropy is therefore a measure of the *errors* in the prediction of the steering angle based on smooth reversal behaviour, and a *higher* value suggests more observed *abrupt* steering wheel adjustments. Steering Entropy (SE) is therefore expected to increase when drivers perform the secondary tasks, and *takes into account* individual differences between drivers.

To summarise, of the measures taken from Table 6.1, speed, time/distance headway, steering wheel Reversal Rate (RR), Entropy (SE) and lane deviation (SDLP) can strongly indicate a driving performance change due to secondary tasks, especially as they are related to road safety. Besides these measurements, the *deviations* of speed and headway were also included in the metrics analysis because they are widely accepted in the evaluation of drivers' performance (Brookhuis & De Waard, 2003; Ma & Kaber, 2005; Tornros & Bolling, 2006; Tornros & Bolling, 2005; Tonros & Bolling, 2005).

Typically, in Car-Following, a *higher* SDLP can be taken as an indicator of steering and driving performance *deterioration*, i.e. more erratic lane keeping or increased steering adjustments, which is usually accompanied by a *higher* RR (RR1.5 and/or RR3) by drivers. However, an *increase* in RR that results in a *similar* or lower SDLP can be an indicator of a coping strategy being employed, i.e. increased effort to obtain more *stable* lane control. This is especially important on busy dual-carriageways, i.e. the Car-Following Scenario, where stable lane keeping has a direct relevance to road safety. Other indicators of coping strategies being employed included *increasing* the time and/or distance headway when dual-tasking compared to baseline driving.

In Free-Driving on a single carriageway, SDLP can be important, but less so than for Car-Following. However, an increase in RR, especially RR3 (large sudden steering wheel turns) in Free-Driving indicates more frequent adjustments to a lack of vehicle control (or driver concentration), and therefore driving performance deterioration.

Given that drivers were also instructed to keep a constant *headway* (in Car-Following) or *speed* (in Free-Driving), a greater *deviation* in these two parameters, i.e. SDHW (standard deviation of headway) in Car-Following or SDSP (standard deviation of speed) in Free-Driving, can also indicate deterioration in driving performance. In addition, provided the SDHW or SDSP remained constant (depending on the Scenario), a *change* to the *value* of either parameter (e.g. a higher headway or a lower speed), indicates a compensation strategy being adopted by drivers (e.g. to increase safety margins or the time required for decision-making). Also, by definition, SE represents the predictability of steering behaviour. Therefore, a high SE is used as an indication of deteriorated and less stable steering behaviour.

6.2.3 Controlling Driving Performance Measurements

When comparing performance differences across the various driver characteristics groups, it is important to look at all the driving performance parameters as a whole, and not focus solely on one measure, because of the potential impact of individual differences in each group. In order to further eliminate individual variability, the same driving performance measurements were collected in *baseline* driving for each subject or driver to use as *control*. Separate averages were calculated for each Scenario using three separate baseline collection periods, i.e. before, during and after the secondary tasks. The baseline average for each driver was then subtracted from the performance measurements while performing the secondary tasks, as recommended by Russell (Russell, 1990), cited by Pastor (Pastor et al., 2006), and these differences named after the original ones, with a suffix of "_Ctrl". These *difference-to-control* values were reported only where they were more material to the absolute ones. In addition, when comparing different characteristics groups (where there are more than two groups) for certain parameters, e.g. SDLP, the incremental changes when performing secondary tasks were compared against a *reference group*, for example Experience Group #3, to determine the *relative* difference between Experience Groups.

Note: Most of the driving performance measurements referred to above are not measured directly by the IV, and the experimenter wrote complex programs to derive most of the values, e.g. for Steering Entropy.

6.3 Pre-Analysis

6.3.1 Data Exploring Analysis

Before the main analysis, data were explored for the normality, homogeneity and existing outliers. Most of the driving performance measurements followed the requirements of the ANOVA Test, therefore most analyses were conducted using the ANOVA Test. When a required assumption was violated, a Non-parameter Test would be applied instead. Some outliers were observed in the DEA processes. They were then either excluded from the analysis, or remained, depending on whether the outliers were *legitimate* or not (see Section 5.4.1). Similar to what was described in Chapter 5, for skewed and un-symmetric parameters, or the ones with outliers, non-linear transformations were conducted for these measurements, and the new variables were named after the original ones, with "Trans_" as prefix. For example, the transformed MNHW was named as Trans_MNHW. Again, for a better explanation of the data, the results for transformed parameters were only reported when the conclusion was different from the original parameters. It was found that the headway data from one of the test runs by Subject # 7 were remarkably higher than those collected in other tests, and appeared as outliers, see Figure 6.2 as the Box Plot for MNHW when drivers performing different types of tasks. This dataset was then scrutinised by examining the distance headway from the video data. The length of the road markings, vehicle type and size in the image were used as reference to estimate the range of the headway. It was finally decided that these outliers were due to technical issues, and should be excluded from further analysis. Other parameters derived from data of the same test run were double-checked by the same method.

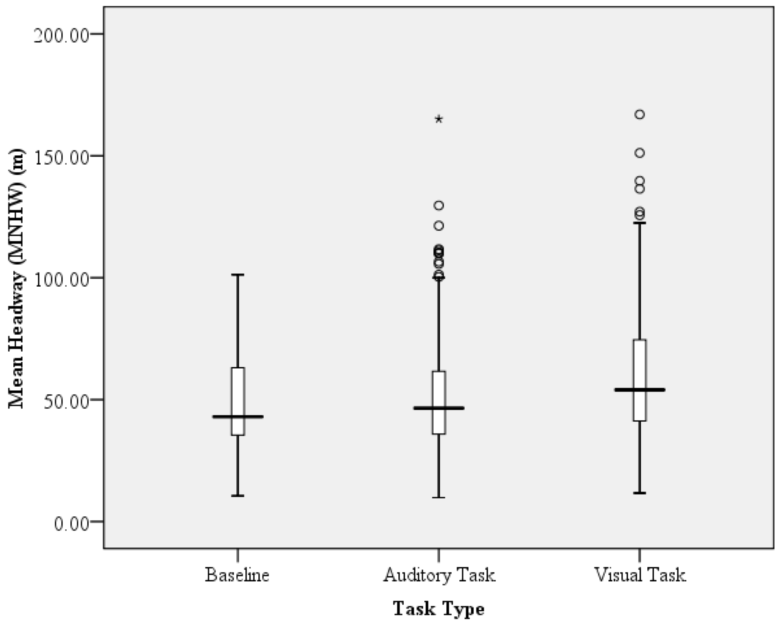


Figure 6.2 Outliers in Mean Headway (MNHW)

6.3.2 Impact of Leading and Tailing Vehicle Types

Previous research has suggested that factors such as types of leading vehicles can cause changes in driving performance (Shehab, 2007). The effects of leading and tailing vehicle types (LVT and TVT respectively) were therefore investigated before the main analysis, to rule out the impact of these factors. Information of leading vehicle and tailing vehicle types were included in the driving performance database, where the leading and tailing vehicle types were differentiated from the video images (collected by the IV during the experiment) according to their size and appearance, and the information of the adjacent vehicle condition (including the leading and tailing vehicle types, and the starting/ending time of each following-episode) were recorded manually (see Chapter 4). As mentioned in Chapter 4, four vehicle types were considered in this research for both LVT and TVT: passenger car (represented as vehicle type "1" in the following analysis), heavy goods vehicle (vehicle type "2" in the following description), van (vehicle type "3"), and no leading or following vehicle within distance which could impact the driver (recorded as "0").

6.3.2.1 Impact of Leading and Tailing Vehicle Types in Car-Following

In the Car-Following scenario, all data were collected when there was a vehicle ahead (i.e. with a leading vehicle), therefore, the LVT investigated were only 1 (Passenger Car as leading vehicle), 2 (Lorry) or 3 (Van).

ANOVA tests showed that LVT had no effect on almost any driving performance measurements, apart from MinTHW ($F(2, 997) = 10.224$, $p < 0.001$). Tukey HSD Post Hoc Tests suggested that the impact on these parameters mainly existed when LVT = "3" (i.e. Van as leading vehicle) comparing with other two groups. The effects of LVT on MinTHW are demonstrated in Table 6.2. In the Car-Following scenario, there were 164 cases when data collected with a passenger car as a leading vehicle (i.e. LVT = 1), 761 cases of a lorry as leading vehicle, and 75 cases of vans as leading vehicles. As cases of following a van ahead was also the minority in the dataset, the cases with van-following were excluded for the analysis in Car-Following.

Table 6.2 Effects of LVT on MinTHW in Car-Following

LVT	Mean	N	SD
Passenger Car	1.20	164	0.28
Lorry	1.19	761	0.31
Van *	1.36	75	0.25

Note: * denotes the p-value < 0.05.

The effect of TVT was investigated in 4 groups: 1: no tailing vehicle; 1: Passenger Car; 2: Lorry; and 3: Van. No effect of TVT has been found for the current dataset. Therefore, in the Car-Following Scenario, cases of all TVT were included in the main analysis.

6.3.2.2 Impact of Leading and Tailing Vehicle Types in Free-Driving

In the Free-Driving Scenario, the purpose of investigating the effect of LVT and TVT was to avoid the effect of leading and tailing vehicle types on the speed choice and other driving behaviour. It was found that LVT had no effect on driving performance measures in any tasking condition ($p < 0.05$). The TVT had impacts on drivers steering behaviour. When there was no task or task demand was relatively low (i.e. performing auditory tasks), the effects mainly reflected on the MNSP and SDSP, while when performing visual tasks, the effect were more noteworthy on steering behaviour ($p < 0.05$). Tukey HSD Post Hoc Tests suggested that the impact on these parameters mainly existed when LVT = "2", i.e. when there was a lorry behind as a tailing vehicle, significantly fewer steering wheel adjustments and lower steering entropy were observed. Again, as there were only 27 cases with a lorry as tailing vehicle (out of 839 cases in total), the cases with lorry as TVT were excluded for the analysis in Free-Driving. The effects of TVT on steering behaviour are demonstrated in Table 6.3.

Table 6.3 Effects of LVT on steering behaviour in Car-Following

TVT	No Tailing Vehicle			Passenger Car			Lorry *			Van		
	Mean	N	SD	Mean	N	SD	Mean	N	SD	Mean	N	SD
RR1.5	34.49	483	15.15	37.28	266	16.56	**28.21**	27	14.90	32.90	63	13.87
SE	0.69	483	0.10	0.70	266	0.10	**0.64**	27	0.14	0.68	63	0.09

Note: * denotes the p-value < 0.05.

6.3.2.3 Summary on the Impact of Leading and Tailing Vehicle Types

The effects of leading and tailing vehicle types (*LVT* and *TVT* respectively) on *driving performance* were investigated, to avoid potential interaction with the impact of secondary tasks and driver characteristics. In the Car-Following scenario, TVT did not affect driving behaviour, but LVT had significant effects on drivers - following a van ahead caused *longer* minimum time headway than for other leading vehicle types. In the Free-Driving Scenario, there were no LVT effects, but TVT did cause some steering behaviour changes - drivers made significantly *less* steering wheel adjustment per minute when there was a lorry behind. Therefore, these cases (where performance was significantly affected by the effects of LVT and TVT) were excluded from the main analysis. In total, 75 cases of following a van were excluded from the Car-Following Scenario, and 27 cases of tailing a lorry excluded from the Free-Driving Scenario, which represents less than 7.5% and 3% of the total respectively.

6.4 Drivers' Individual Difference in the Baseline Driving

Before analysing the impact of the secondary task on the driving performance, some background information were required on how different drivers' characters (e.g. gender, age/experience etc.) affect the driving performance in baseline driving, i.e., the normal driving without doing any additional secondary task. The One-Way ANOVA indicates that the driving performance varies enormously between the two driving scenarios – Car-Following and Free-Driving. Therefore the analysis was performed in two sets: Car-Following and Free-Driving.

6.4.1 Baseline Driving in Car-Following Scenario

6.4.1.1 Gender Difference

In baseline driving from Car-Following Scenario, some gender differences were found, mainly existing in speed and headway choices. Female drivers kept a slightly lower speed, a shorter

MNHW, a shorter MinTHW,and lower SDTHW ($p< 0.05$) in baseline driving, see Table 6.4. The “Diff (%)” column shows the difference between female and male, in percentage, which was calculated by the following formula:

$$\text{Gender Diff} = \frac{\text{mean(Female)} - \text{mean(Male)}}{\text{mean(Male)}} \tag{6.6}$$

Table 6.4 Gender difference in driving performance in baseline Car-Following

Gender		MNSP*	SDSP	RR1.5	RR3	SE	MNHW*	SDTHW*	MinHW*
Male	Mean	**85.46**	2.32	20.32	7.42	0.56	**51.18**	**0.67**	**44.32**
	SD	5.96	1.94	10.83	7.98	0.12	22.03	0.37	20.74
Female	Mean	**82.44**	1.76	21.58	6.69	0.56	**42.66**	**0.52**	**37.29**
	SD	4.89	1.27	11.41	7.04	0.12	14.72	0.17	13.83
Diff (%)	-3.53	-24.14	6.20	-9.84	0.00	-16.65	-22.39	-15.86	-3.53

Note: * indicates a significant difference existing between the three groups.

6.4.1.2 Experience and Age

No significant difference of speed choice was found between three experience groups. The results from ANOVA revealed that measurements of lane keeping, Reversal Rate, and RR3, SE, and Headway parameters (MNHW, MNTHW, ,MinHW, and MinTHW) by experience groups is statistically significant ($p < 0.05$). The Tukey Test for the Post Hoc analysis provided some more detailed information:

1. Experience Group #1 had a remarkably higher SDLP (indicating worse lane control), more per minute steering wheel adjustments (both RR1.5 and RR3), higher SE (indicating a less predictability in steering behaviour), and a slightly lower MinTHW, comparing with Experience Groups #2 and #3. It is noteworthy that they made more steering adjustments, but led to a worse lane control, also, slightly shorter headways.
2. Experience Group 2 kept a much longer distance and time headway. Drivers from this group also kept the longest minimum distance and time head way.
3. Except SDHW and SDTHW, all headway measurements were sensitive to experience.

Table 6.5 illustrates the difference on driving performance between the three experience groups. It may be seen that the least experienced drivers showed less *capability* than other drivers, even in baseline driving.

Table 6.5 Effect of experience on driving performance in baseline Car-Following

Experience Group	#1		#2		#3		Diff 1 & 3
	Mean	SD	Mean	SD	Mean	SD	
MNSP (kph)	84.48	4.58	84.35	5.68	84.31	7.09	0%
SDSP (kph)	2.33	1.76	1.86	1.34	2.04	1.95	14%
SDLP (m)*	**0.18**	0.08	0.14	0.08	0.13	0.07	33%
RR1.5 (times/min)*	**25.13**	12.42	17.95	8.54	17.3	8.74	45%
RR3 (times/min)*	**10.16**	9.76	4.83	4.58	5.03	4.54	102%
SE*	**0.59**	0.13	0.54	0.1	0.53	0.11	12%
MNHW (m)*	42.32	17.38	**61.25**	19.81	47.46	20.06	-11%
MNTHW (s)*	1.94	0.31	**2.14**	0.27	2	0.2	-4%
MinHW (s)*	36.23	15.96	**54.3**	19.41	41.22	18.49	-12%
MinTHW (s)*	**1.09**	0.29	**1.37**	0.17	1.26	0.25	-14%

Note: * indicates a significant difference existing between the three groups.
* Diff 1 & 3 shows the difference between Group #1 and Group #3 in percentage calculated by: (Value for #Group1 - Value for #Group 3)/ Value for #Group1.

The differences between age groups had a similar trend with that of the experience group, i.e., older drivers (Group #3) shows less minor steering wheel turns, while the younger drivers showed the highest steering entropy ($p < 0.05$); Group #2 kept a longer distance and time headway, but this effect was not significant. The effect of age is less noticeable than that of experience. This may suggest that for the current sample, the experience factor has a higher impact than that of age.

6.4.1.3 Annual Mileage

Drivers with the least annual mileage showed significantly different driving performance. The ANOVA and Tukey Post Hoc Tests ($p < 0.05$) suggested that less annual mileage (Annual Mileage Group #1) showed a remarkably higher lane deviation, more large steering wheel turns, but smaller headway, MinHW, MinTHW than that of Group #2 and Group #3 ($p < 0.05$). Similar to the effect of Experience, the group with the least annual mileage showed less *capability* than other drivers, even in baseline driving. For more detail, see Table 6.6.

Table 6.6 Effect of annual mileage on driving performance in baseline Car-Following

Mileage Group	# 1		# 2		# 3	
	Mean	SD	Mean	SD	Mean	SD
MNSP	84.14	5.58	84.64	4.43	84.38	6.89
SDSP	2.15	1.65	2.39	1.76	1.87	1.80
SDLP *	**0.19**	0.08	0.14	0.08	0.13	0.07
RR1.5	22.25	11.28	22.51	11.62	18.15	9.94
RR3 *	**10.09**	8.54	6.55	7.91	5.38	5.95
SE	0.57	0.11	0.57	0.14	0.55	0.11
MNHW *	**37.08**	19.90	55.66	17.33	50.88	18.72
MNTHW	1.84	0.25	2.16	0.27	2.01	0.24
MinHW *	**31.50**	17.72	47.60	16.12	45.32	18.52
MinTHW *	**0.98**	0.28	1.31	0.18	1.30	0.23

6.4.2 Baseline Driving in Free-Driving Scenario

As the Free-Driving Scenario was designed to create a condition where drivers could control their speed without interacting with leading vehicles, all data were collected when there was no leading vehicle which could impact the speed choice of the IV. Hence the headway data were not meaningful in this scenario. Only the MNSP, SDSP, SDLP, RR1.5, RR3, and SE data were analyzed and reported in the Free-Driving.

6.4.2.1 Gender Difference

For baseline driving in Free-Driving, female drivers showed a lower SDSP, lower RR1.5 (both $p < 0.05$) and marginally lower SE ($p = 0.058$), suggesting a more stable vehicle control, see Table 6.7.

Table 6.7 Gender difference in driving performance in baseline Free-Driving

	Male		Female		df	F	p_value
	Mean	SD	Mean	SD			
MNSP	82.09	5.52	82.30	3.55	1,85	0.05	0.83
SDSP*	8.10	3.80	6.25	3.50	1,85	5.45	0.02
SDLP	0.18	0.08	0.18	0.08	1,85	0.03	0.86
RR1.5*	30.32	11.60	25.66	17.34	1,85	2.25	0.14
RR3	9.59	7.38	8.97	10.11	1,85	0.11	0.74
SE*	0.66	0.09	0.62	0.12	1,85	3.69	0.06

6.4.2.2 Experience/Age and Mileage

Steering behaviour was significantly different between three experience groups (RR1.5, RR3, $p < 0.05$). Post Hoc analysis (Turkey HSD Test) confirmed that the most experienced drivers had a lower number of steering wheel adjustments per minute (RR1.5 and RR3), and a better predictability of steering wheel adjustments (lower SE). For more details, see Table 6.8.

Table 6.8 Effect of experience on driving performance in baseline Free-Driving

Experience Group	# 1		# 2		# 3		Diff 1 & 3
	Mean	SD	Mean	SD	Mean	SD	
MNSP	81.17	3.4	81.36	3.58	83.88	6.16	-3%
SDSP	7.76	3.12	7.58	3.44	6.57	4.57	15%
SDLP	0.19	0.1	0.18	0.07	0.18	0.08	5%
RR1.5 *	33.1	17.55	29.42	9.16	**22.09**	11.37	33%
RR3 *	12.68	10.75	8.5	4.65	**6.09**	6.61	52%
SE *	0.66	0.12	0.67	0.09	**0.61**	0.08	8%

Note: * indicates a significant difference existing between the three groups.
* Diff 1 & 3 shows the difference between Group #1 and Group #3 in percentage calculated by: (Value for #Group1 - Value for #Group 3)/ Value for #Group1.

Similar effects on steering behaviour were found between the three age groups. In detail, Age Group #3 had a lower RR1.5, lower RR3 and a lower SE (ANOVA, $p < 0.05$ and Tukey Post Hoc Tests, $p < 0.05$).

The drivers' annual mileage had significant effect on some measurements for driving performance in Free-Driving Scenario. SDSP, SDLP, RR1.5 and RR3 were found to be different between three mileage groups ($p < 0.05$). Post Hoc analysis confirmed that drivers with the least annual mileage (Annual Mileage Group #1) had a much higher lane deviation, while the most experienced group (Annual Mileage Group #3) showed lowest SDSP, and steering wheel turns (RR1.5 and RR3).

6.4.3 Summary for Driving Performance in Baseline Driving

In the Car-Following Scenario, female drivers kept a lower speed, shorter headway and minimum time headway, and a lower deviation in time headway. Drivers with less experience (driving experience less than 10 years) showed a remarkably larger deviation in lane position than the other two groups, more adjustments of steering behaviour, suggesting a worse lateral control by the less experienced drivers. This was also confirmed by higher steering entropy from this group, which indicates less predictability of steering behaviour. Drivers with "medium experience" (driving experience from 11 to 20 years) kept significantly longer headway (and minimum headway) and time headway (and minimum time headway). The effect of age showed a similar tendency, but was less significant. Results from mileage groups confirmed the effect of experience, where the least annual mileage group showed a remarkably higher lane deviation, more large steering (RR3) wheel adjustments (suggesting a more abrupt steering turns); and showed the relatively smaller headway, despite the poor steering performance.

In the Free-Driving Scenario, female drivers had a lower speed deviation, lower small wheel adjustments and marginally lower steering entropy, suggesting a more stable vehicle control. Experience, age, and annual mileage showed similar tendencies, i.e. the most experienced, higher age and higher annual mileage group made less steering wheel adjustments and lower steering entropy; while the younger drivers who have less experience showed a much higher steering wheel adjustments and less predictability in their steering. Especially, the lowest annual mileage group showed a significantly higher deviation in lane keeping.

It was therefore concluded that in both driving scenarios, female drivers showed a more stable vehicle control (i.e. lower headway deviation in Car-Following, and lower speed deviation in Free-Driving), compared to male drivers. And less experience/age groups showed higher steering wheel adjustments, their steering wheel behaviour was less predictable, and they had higher lane deviation, compared to more experienced/higher age drivers.

6.5 Impact of Secondary Tasks on the Driving Performance

6.5.1 Performing Auditory Tasks in Car-Following

6.5.1.1 Task Level

Driving performance in 4 levels - three task difficulty levels and baseline – was compared and analysed. It was found that task level had significant impact on RR1.5 ($F (3, 538) = 4.224$, $p = 0.006$) and SE ($F (3, 538) = 5.509$, $p = 0.001$). Post Hoc analysis suggested that all differences existed only between baseline and auditory tasking, and no statistically significant difference was found between 3 task difficulty levels (Turkey HSD Tests, $p < 0.05$). Performing the auditory tasks when following raised minor steering wheel adjustments and caused higher steering entropy (indicating less predictable steering behaviour) compared to the baseline driving. As no difference was found between the three task difficulty levels, the data when performing auditory tasks were combined together to form a bigger but still homogeneous dataset – naming “Auditory Tasking”. Comparison was then conducted between baseline and tasking condition. It was found further, the deviation of speed increased when performing auditory tasks (Baseline: 2.12; Tasking: 2.52, $F (1, 588) = 4.450$, $p = 0.035$). Detailed driving performance measurements of when drivers were engaged in the auditory tasks while Car Following (compared to baseline) are shown in Table 6.9. The column named “Tasking” presents the mean values of the whole dataset when performing auditory tasks in each Scenario. The “Diff” column shows the performance changes from baseline to when performing tasks, in percentage, which was calculated by the following formula:

$$\text{Diff} = \frac{\text{mean(Tasking)} - \text{mean(Baseline)}}{\text{mean(Baseline)}} \qquad (6.7)$$

Table 6.9 Driving performance between baseline and auditory tasking in Car-Following

Task Level	Baseline	Level 1	Level 2	Level 3	Tasking	Diff (%)
MNSP	84.39	84.57	84.89	85.48	85.00	1%
SDSP *	**2.12**	2.45	2.55	2.55	**2.52**	**19%**
SDLP	0.15	0.14	0.14	0.15	0.15	-5%
RR1.5 *	**20.77**	24.26	25.44	24.97	**24.92**	**20%**
RR3	7.16	8.05	7.50	8.27	7.94	11%
SE *	**0.56**	0.60	0.60	0.61	**0.61**	**9%**
MNHW	48.08	49.95	48.77	51.15	49.97	4%
SDHW	4.50	5.15	5.14	5.08	5.12	14%
MNTHW	2.00	2.04	2.02	2.04	2.03	1%
SDTHW	0.62	0.67	0.65	0.66	0.66	7%
MinHW	41.76	42.21	41.10	43.67	42.34	1%
MinTHW	1.21	1.17	1.15	1.18	1.17	-3%

6.5.1.2 Gender

The gender difference was more noteworthy while performing auditory tasks. Performing auditory tasking did not affect any driving measurements for male drivers, compared with baseline driving (ANOVA and Mann-Whitney U Test, $p > 0.05$), apart from an increase in speed deviation (Trans_SDSP, $F(2, 348) = 4.625$, $p = 0.032$). While for females, performing the auditory tasks caused increases in speed ($F(3, 190) = 7.475$, $p = 0.007$) and SDSP ($F(3, 190) = 4.035$, $p = 0.046$), a higher steering wheel adjustments (RR1.5, ANOVA and Mann-Whitney U Test, $p < 0.05$); RR3, ANOVA and Mann-Whitney U Test, $p<0.05$), less predictable steering behaviour (SE: ($F(3,190) = 15.339$, $p < 0.001$) and MNHW ($F(3, 188) = 3.536$, $p = 0.016$). Tukey Post Hoc analysis showed that these effects only existed between baseline and when performing the auditory tasks, but not between 3 levels of task difficulty ($p > 0.05$). Table 6.10 presents the comparison between male and females' driving performance while performing auditory tasks and in baseline.

When performing the auditory tasks, compared to the male drivers, females had lower deviation in speed (Trans_SDSP, $F(1, 396) = 3.718$, $p = 0.021$), more steering wheel turns (RR1.5, Mann-Whitney U Test, $z = -5.056$, $p < 0.001$; RR3, $z = -3.081$, $p < 0.001$), kept lower deviation in time headway (SDTHW, $z = -2.337$, $p = 0.019$), a smaller SDLP ($F(1, 396) = 3.951$, $p = 0.048$), a lower SDHW ($F(1, 396) = 3.953$, $p = 0.047$) and marginal longer minimum distance headway ($F(1,383) = 3.039$, $p = 0.082$).

Table 6.10 Gender difference in driving performance in ATCF

Gender	Male			Female			Gender Diff in Baseline	Gender Diff in AT
	BL	AT	Diff	BL	AT	Diff		
MNSP	85.46	85.05	0%	**82.44**	**84.91**	**3%**	**-4%**	0%
SDSP*	**2.32**	**2.64**	**14%**	**1.76**	**2.30**	**30%**	-24%	**-13%**
SDLP*	0.156	0.151	-4%	0.145	0.136	-6%	-7%	**-10%**
RR1.5*	20.32	22.42	10%	**21.58**	**29.46**	**37%**	6%	**31%**
RR3*	7.42	6.94	-6%	**6.69**	**9.77**	**46%**	-10%	**41%**
SE	0.56	0.59	6%	**0.56**	**0.64**	**14%**	0%	8%
MNHW	51.18	49.14	-4%	**42.66**	**51.40**	**20%**	**-17%**	5%
SDHW*	4.86	5.57	15%	3.88	4.35	12%	-20%	**-22%**
MNTHW	2.01	2.03	1%	1.98	2.03	2%	-1%	0%
SDTHW*	0.67	0.71	5%	0.52	0.58	11%	**-22%**	**-18%**
MinHW*	44.32	41.01	-7%	**37.29**	**44.65**	**20%**	-16%	**9%**
MinTHW	1.19	1.16	-3%	1.23	1.18	-4%	4%	2%

Note: * indicates a significant difference existing between genders when performing auditory tasks.

As explained in Section 6.4.1.1, there existed a gender difference in the baseline driving in headway choice (i.e. female had a lower headway). However, this difference vanished when performing auditory tasks. To investigate the reason, the interactive effect of headway between the gender and task condition has been studied. The result suggests that there is an interaction effect for the speed with the task level and gender ($F (3,527) = 3.234$, $p = 0.022$), suggesting, females increased headway when performing auditory tasks, while male drivers did not change their headway decision. The increased speed for female drivers was very small, and in a Car-Following Scenario, this could be merely due to the speed choice of the vehicles.

Generally, the auditory task condition changed the driving behaviour more for female drivers than for males. For example, there was a significant increase in the per-minute steering wheel turns for female, but not for male drivers. This explained why even though no difference in the steering behaviour for the two genders in baseline driving was found, there existed a difference when performing secondary task.

This suggests that female drivers adopt their steering behaviour change when performing the auditory tasks to compensate. The compensation strategy used by female drivers was also evidenced by the significant increase in headway during auditory tasking. This strategy also led to a more stable lateral and longitudinal control for female drivers, i.e. a lower SDLP, lower SDSP, lower SDHW and lower SDTHW when performing the auditory tasks. They have made more effort in steering, by an increased number of steering wheel adjustments, to keep a lower SDLP. This may reflect a more conservative driving strategy or over-compensation.

6.5.1.3 Experience/Age and Mileage

When performing auditory tasks in Car-Following, Experience Group #2 and #3 increased SDSP and steering adjustments (RR1.5, RR3), comparing with baseline driving, but Group #1 did not. But even after this, they still kept the same level of SDSP with the Group #1, and lower reversal rate than Group #1.

When comparing the driving performance between three experience groups when performing auditory tasks, the tendency of experience difference was similar to the baseline. Experience Group #1 still kept a higher SDLP ($F (2, 395) = 3.322$, $p = 0.037$), more steering wheel turns (RR1.5, $F (2, 395) = 10.475$, $p < 0.001$; RR3, $F (2, 395) = 13.556$, $p < 0.001$) and higher Steering Entropy ($F (2, 395) = 6.763$, $p < 0.001$). Despite the poor steering and lane keeping, when performing auditory tasks, this group of drivers showed the shortest headway (MNHW, $F (2, 395) = 4.343$, $p = 0.014$; MinHW, $F (2, 395) = 5.118$, $p = 0.006$; MinTHW, $F (2, 395) = 9.863$, $p < 0.001$). Table 6.11 shows the driving performance by experience group, when performing auditory tasks in Car-Following. These evidences suggest that less experienced drivers (driving experience less than 10 years) had a significantly worse lane keeping and steering behaviour both in baseline and when performing the auditory tasks. Compared with more experienced drivers, they did not compensate enough, for example, they showed a significant shorter headway in auditory tasking.

Table 6.11 Effect of experience on driving performance in ATCF

Experience Group	# 1			# 2			# 3			AT Diff 1 & 3
Task Condition	BL	AT	Diff	BL	AT	Diff	BL	AT	Diff	
MNSP	84.48	85.14	1%	84.35	85.14	1%	84.31	84.56	0%	1%
SDSP	2.33	2.41	4%	**1.86**	**2.51**	**35%**	**2.04**	**2.74**	**34%**	-12%
SDLP*	**0.18**	**0.15**	-12%	0.14	0.13	-3%	0.13	0.14	7%	**10%**
RR1.5*	**25.13**	**27.89**	11%	**17.95**	**22.55**	**26%**	**17.3**	**21.83**	**26%**	**28%**
RR3*	**10.16**	**9.9**	-3%	**4.83**	**5.71**	**18%**	**5.03**	**6.72**	**34%**	**47%**
SE*	**0.59**	**0.63**	6%	0.54	0.6	10%	0.53	0.57	9%	**10%**
MNHW*	42.32	**46.65**	10%	**61.25**	52.58	-14%	47.46	53.8	13%	**-13%**
SDHW	4.82	5.29	10%	5.39	5.01	-7%	3.6	4.9	36%	8%
MNTHW	1.94	2.02	4%	**2.14**	2.05	-4%	2	2.04	2%	-1%
SDTHW	0.66	0.72	8%	0.59	0.55	-7%	0.57	0.66	16%	9%
MinHW*	36.23	**39.22**	8%	**54.3**	44.44	-18%	41.22	46.34	12%	**-15%**
MinTHW*	**1.09**	**1.11**	2%	**1.37**	1.29	-6%	1.26	1.14	-9%	**-3%**

Note: * indicates a significant difference existing between experience groups when performing auditory tasks.
* AT Diff 1 & 3 shows the difference between Group #1 and Group #3.

The effect of auditory tasks on the different age groups enlarged the age effects. When performing auditory tasks, the RR1.5 ($Chi^2(2)=15.393$, $p<0.001$), SE ($Chi^2(2)=27.779$, $p<0.001$), MNHW ($Chi^2(2)=8.988$, $p=0.011$), MinHW ($Chi^2(2)=8.792$, $p=0.012$) and MinTHW ($Chi^2(2)=24.567$, $p<0.001$) showed significant differences between the three age groups. Age Group #1 did more steering wheel turns, showed the largest SE, while Age Group #2 kept the longest headway, largest minimum distance and time headway. When considering the behaviour changes between baseline and tasking of each age group: Age Group #1, who used to have the highest RR1.5 in baseline, increases RR1.5 even more significantly ($F(1,182)=4.316$, $p=0.039$), and increased SE ($F(1,182)=6.190$, $p=0.014$); Group #3 increased SE in task performing compared to baseline ($F(1,163)=3.912$, $p=0.050$). There was no performance change for the Group #2 when performing the auditory tasks. This may due to their longer headway. Table 6.12 illustrates the comparison of driving performance between three age groups.

Table 6.12 Effect of age on driving performance in ATCF

Age Group	# 1			# 2			# 3		
Task Condition	BL	AT	Diff	BL	AT	Diff	BL	AT	Diff
MNSP	84.70	85.23	1%	84.76	85.36	1%	83.85	84.32	1%
SDSP	2.33	2.45	5%	2.03	2.52	24%	2.01	2.60	29%
SDLP	0.16	0.15	-6%	0.15	0.14	-11%	0.14	0.15	3%
RR1.5*	23.56	**28.98**	23%	21.94	23.80	8%	**17.48**	21.57	23%
RR3	7.66	9.76	27%	8.87	6.96	-21%	5.42	7.00	29%
SE*	**0.59**	**0.65**	10%	0.57	0.60	5%	0.53	0.57	8%
MNHW	48.36	50.14	4%	51.43	**52.14**	1%	45.51	47.42	4%
SDHW	4.93	5.61	14%	5.20	5.08	-2%	3.65	4.61	26%
MNTHW	1.99	2.02	2%	2.05	2.05	0%	1.98	2.01	2%
SDTHW	0.71	0.68	-5%	0.57	0.56	-2%	0.56	0.74	32%
MinHW	42.56	42.41	0%	44.57	**44.04**	-1%	39.12	40.43	3%
MinTHW	1.19	1.14	-4%	1.23	**1.30**	5%	1.20	1.05	-12%

Note: * indicates a significant difference existing between age groups when performing auditory tasks.

The effects of performing auditory tasks on annual mileage groups were exactly the same as that for the experience groups. No significant behaviour change was found for Group #1, while Groups #2 and #3 were found to compensate their driving performance for the task demand. When comparing the driving performance between the three groups when performing auditory tasks, Mileage Group #1 showed higher deviation in lane position (SDLP: ANOVA, $F(2, 395) = 9.226$, $p < 0.001$), and higher large steering wheel adjustments (RR3: ANOVA, $F(2, 395) = 4.435$, $p = 0.012$). Again, this group also showed shorter headway and minimum headway. Table 6.13 gives a detailed comparison of driving performance between the three mileage groups.

Table 6.13 Effect of mileage on driving performance in ATCF

Mileage Group	# 1			# 2			# 3		
Task Condition	BL	AT	Diff	BL	AT	Diff	BL	AT	Diff
MNSP	84.14	85.65	2%	84.64	84.92	0%	84.38	84.55	0%
SDSP	2.15	2.65	23%	2.39	2.46	3%	**1.87**	**2.48**	**32%**
SDLP*	**0.19**	**0.17**	-13%	0.14	0.14	0%	0.13	0.13	2%
RR1.5	22.25	24.01	8%	22.51	26.01	16%	**18.15**	**24.35**	**34%**
RR3*	**10.09**	**9.63**	-5%	6.55	7.68	17%	**5.38**	**6.80**	**27%**
SE	0.57	0.58	3%	**0.57**	**0.62**	**10%**	**0.55**	**0.61**	**11%**
MNHW*	**37.08**	**43.24**	17%	55.66	54.39	-2%	50.88	50.75	0%
SDHW	5.00	5.25	5%	5.74	5.11	-11%	**3.16**	**5.03**	**59%**
MNTHW*	1.84	**1.90**	3%	2.16	2.13	-1%	2.01	2.03	1%
SDTHW	0.61	0.56	-9%	0.71	0.74	4%	0.54	0.65	19%
MinHW*	**31.50**	**35.46**	13%	47.60	46.62	-2%	45.32	43.44	-4%
MinTHW*	**0.98**	**0.96**	-2%	1.31	1.30	-1%	**1.30**	**1.19**	**-8%**

Note: * indicates a significant difference existing between mileage groups in ATCF.

6.5.1.4 Experience of and Attitude towards Auditory IVISs

The experience of previous using auditory systems had no effect on driving performance. Some interesting findings were revealed for drivers who considered using auditory IVISs as safe – this group showed a shorter headway than the other two (MNHW, $F(3, 382) = 5.114$, $p = 0.006$; MNTHW, $F(3, 382) = 14.154$, $p < 0.001$; MinTHW, $F(3, 382) = 14.778$, $p < 0.001$). These results answered the question why they also had a higher workload level, found in Chapter 5. The choice of relatively short headway has increased the whole task demands for this group. The effect of drivers' attitude towards auditory tasks is listed in Table 6.14.

Table 6.14 The effect of attitude towards auditory systems on driving performance in ATCF

Atti_Audi Group	# 1		# 2		# 3		AT Diff 1 & 3
Task Condition	BL	AT	BL	AT	BL	AT	
MNHW	43.5	**44.63**	40.98	50.06	59.99	53.73	-20%
MNTHW	1.84	**1.91**	2.01	2.08	2.1	2.07	-8%
MinTHW	1.05	**1.03**	1.17	1.17	1.36	1.26	-22%

Note: * indicates a significant difference existing between Attitude Groups when performing auditory tasks.
* AT Diff 1 & 3 shows the difference between Group #1 and Group #3 when performing auditory tasks, calculated by: (Value for #Group1 - Value for #Group 3)/ Value for #Group1.

6.5.1.5 Repetition of the Task Performing

The effect that repetition of the secondary task performing had on the driving performance was investigated, to understand whether driving performance improved while drivers became more familiar with the secondary tasks. No positive result has been found, which suggests driving performance was not affected by the repetition of task performing.

6.5.2 Performing Visual Tasks in Car-Following

6.5.2.1 Task Level Effect

The difficulty level of the visual tasks impacted significantly on the following measurements of driving performance: SDLP ($F(3, 549) = 8.036$, $p < 0.001$), RR1.5 ($F(3, 549) = 15.079$, $p < 0.001$), RR3 ($Chi^2(3) = 53.8$, $p < 0.001$), steering entropy ($F(3, 549) = 16.529$, $p < 0.001$), MNHW ($F(3, 536) = 8.789$, $p < 0.001$), MNTHW ($F(3, 536) = 4.731$, $p = 0.003$), and MinHW ($F(3, 536) = 7.334$, $p < 0.001$). The Tukey and Games-Howell Post Hoc analysis showed that the task-level effects mainly exist between with and without task performing. The differences between the three task difficulty levels were mostly not significant. The detailed results are illustrated in Table 6.15.

Table 6.15 Driving performance between baseline and visual tasking in Car-Following

Task Level	Baseline	Level 1	Level 2	Level 3	Tasking	Diff
MNSP	84.39	83.66	84.41	83.16	83.73	-1%
SDSP	2.12	2.25	2.53	2.47	2.4	14%
SDLP*	**0.15**	0.17	0.18	0.20	**0.18**	**18%**
RR1.5*	**20.77**	27.75	28.63	28.23	**28.19**	**36%**
RR3*	**7.16**	12.14	13.49	14.23	**13.28**	**85%**
SE*	**0.56**	0.62	0.63	0.64	**0.63**	**13%**
MNHW*	**48.08**	58.08	61.58	59.21	**59.58**	**24%**
SDHW	4.50	5.59	5.61	5.57	5.59	24%
MNTHW*	**2.00**	2.09	2.11	2.12	**2.11**	**5%**
SDTHW	0.62	0.68	0.76	0.72	0.72	17%
MinHW*	**41.76**	49.41	53.23	50.93	**51.14**	**22%**
MinTHW	1.21	1.23	1.21	1.21	1.22	1%

Note: * indicates a significant difference existing between baseline and when performing visual tasks.

6.5.2.2 Gender Difference

Comparing the gender difference by investigating the driving performance change from the baseline to visual tasking condition, results showed that for both genders, there were significant increases in the SDLP and headway. Males and females had a change in RR1.5, RR3 and Entropy to almost the same level, from baseline driving to performing visual tasks.

Performing visual tasks caused more gender differences in driving performance. Apart from females' lower speed ($z = -5.431$, $p < 0.001$), the same as that in baseline, SDLP, MNHW, SDHW, SDTHW and MinHW were all significantly different between male and female drivers when performing visual tasks. The female drivers were found to keep a smaller SDLP, which suggested a better lane keeping ($Z = -2.714$, $p = 0.007$). Shorter headway (MNHW, $z = -2.854$, $p = 0.004$) by (11.91%), and a shorter minimum headway (MinHW, $z = -2.706$, $p = 0.007$) were still found for female drivers. Also, less variation in both distance and time headway was found for female drivers in tasking. Table 6.16 gives the detailed results.

It can be concluded that performing visual tasks had an impact on both genders, but female drivers managed to keep a lower deviation in lane keeping, and kept a more stable headway by compromise of speed (i.e. driving in a lower speed). This improved primary driving performance may have a trade-off from the "worse" secondary task performance (Section 5.5.2). Combined with the information that females rate the task higher in the subjective rating (Section 5.4.3), female drivers probably have made more efforts to maintain an "unaffected" driving performance when performing visual tasks.

Table 6.16 Gender difference in driving performance in VTCF

Gender	Male			Female			Gender Diff in Baseline	Gender Diff in Tasking
	BL	VT	Diff	BL	VT	Diff		
MNSP*	85.46	84.67	-1%	82.44	81.76	-1%	**-4%**	**-3%**
SDSP	2.32	2.46	6%	**1.76**	**2.32**	**31%**	-24%	-6%
SDLP*	**0.16**	**0.19**	**19%**	**0.14**	**0.17**	**17%**	-7%	**-8%**
RR1.5*	**20.32**	**26.91**	**32%**	**21.58**	**30.89**	**43%**	6%	**15%**
RR3*	**7.42**	**12.25**	**65%**	**6.69**	**15.43**	**131%**	-10%	**26%**
SE*	**0.56**	**0.62**	**11%**	**0.56**	**0.65**	**16%**	0.4%	**4%**
MNHW*	**51.18**	**62.12**	**21%**	**42.66**	**54.43**	**28%**	**-17%**	**-12%**
SDHW*	**4.86**	**6.02**	**24%**	3.88	4.71	21%	-20%	**-22%**
MNTHW	2.01	2.11	5%	1.98	2.10	6%	-1.4%	-0. 5%
SDTHW*	0.67	0.77	15%	0.52	0.61	17%	**-22%**	**-21%**
MinHW*	**44.32**	**53.21**	**20%**	**37.29**	**46.94**	**26%**	**-16%**	**-12%**
MinTHW	1.19	1.22	2.2%	1.23	1.21	-1.5%	4%	-0%

Note: * indicates a significant difference existing between genders when performing visual tasks.

* Diff_Gender was calculated by: (Measured value for female – value for male) / value for male; while Diff. Baseline & Tasking is calculated by: (Measured value when performing tasks – value for baseline) / value for baseline.

6.5.2.3 Experience/Age and Mileage

For all three experience groups, driving performance changed to same extent while performing the visual tasks. Performing the visual tasks while driving did not change the tendencies that appeared in baseline driving. However, as drivers in all groups adapted their primary driving to the dual-tasking situation, it became even more clear that drivers with the least driving experience performed the highest number of steering wheel adjustments, had the highest steering entropy and deviation in lane position, but also kept shorter headway, both in terms of distance and time.

Steering entropy showed to be a very sensitive measurement for the steering behaviour change. In task performing, SE for all experience groups increased compared to baseline. The difference between the experience groups became more noteworthy, i.e. the less the experience, the higher the SE.

SDLP, RR1.5, RR3 and steering entropy increased in all three experience groups when performing the visual tasks, compared to the baseline driving. For Groups #1 and #3, the increases in the mean and minimum distance headway, and mean time headway were also significant. These changes, however, were not found in Experience Group #2. In short, performing the visual tasks changed the driving performance of drivers in Group #2 the least and Group #1 the most. The detailed results are shown in Table 6.17.

Table 6.17 Effect of experience on driving performance in VTCF

	BL			VT			Diff (Baseline & Tasking)			VT Diff 1&3
Experience	1	2	3	1	2	3	1	2	3	
MNSP	84.48	84.35	84.31	83.28	85.4	82.73	-1%	1%	-2%	1%
SDSP	2.33	1.86	2.04	2.57	2.25	2.32	10%	21%	14%	10%
SDLP	**0.18**	0.14	0.13	**0.2**	0.16	0.17	**11%**	**18%**	**31%**	15%
RR1.5	**25.13**	17.95	17.3	**30.05**	26.66	26.69	**20%**	**49%**	**54%**	11%
RR3	**10.16**	4.83	5.03	**15.73**	10.76	11.79	**55%**	**123%**	**135%**	25%
SE	**0.59**	0.54	0.53	**0.65**	0.62	0.61	**10%**	**14%**	**16%**	6%
MNHW	42.32	**61.25**	47.46	**53.23**	67.12	62.93	**26%**	10%	**33%**	-18%
SDHW	4.82	5.39	3.6	5.54	5.54	5.72	15%	3%	59%	-3%
MNTHW	1.94	**2.14**	2	2.09	2.1	2.13	**8%**	-2%	**6%**	-2%
SDTHW	0.66	0.59	0.57	**0.78**	0.63	0.7	17%	5%	24%	10%
MinHW	36.23	**54.3**	41.22	**44.51**	58.99	54.67	**23%**	9%	**33%**	-23%
MinTHW	1.09	**1.37**	1.26	1.14	1.31	1.25	5%	-4%	-1%	-10%

Note: * indicates a significant difference existing between Experience Groups when performing visual tasks.

* VT Diff 1 & 3 shows the difference between Group #1 and Group #3 when performing visual tasks, calculated by: (Value for #Group1 - Value for #Group 3)/ Value for #Group1.

While performing visual tasks, the effect of age became significant. Age Group #1 showed a much larger SDLP ($F(2, 443)=5.182$, $p=0.006$), RR1.5 ($F(2, 443)=11.586$, $p<0.001$), RR3 ($F(2, 443)=8.500$, $p<0.001$) and SE ($Chi^2(2)=39.196$, $p<0.001$), while Age Group #2 kept a longer minimum time headway ($Chi^2(2)=20.044$, $p<0.001$), with lower standard deviation (SDTHW, $Chi^2(2)=11.704$, $p=0.003$). The details are shown in Table 6.18.

Table 6.18 Effect of age on driving performance in VTCF

	BL			VT			Diff (Baseline & VT)		
Age Group	1	2	3	1	2	3	1	2	3
MNSP	84.70	84.76	83.85	84.55	83.61	83.05	0%	-1%	-1%
SDSP	2.33	2.03	2.01	2.59	2.24	2.41	11%	11%	**20%**
SDLP*	0.16	0.15	0.14	**0.19**	0.17	0.18	**20%**	9%	**26%**
RR1.5*	23.56	21.94	**17.48**	**32.10**	26.21	26.37	**36%**	**19%**	**51%**
RR3*	7.66	8.87	5.42	**16.06**	11.37	12.48	**110%**	28%	**130%**
SE*	**0.59**	0.57	0.53	**0.68**	0.61	0.61	**15%**	**7%**	**16%**
MNHW	48.36	51.43	45.51	61.47	57.14	60.04	**27%**	11%	**32%**
SDHW	4.93	5.20	3.65	6.07	4.98	5.69	23%	-4%	**56%**
MNTHW	1.99	2.05	1.98	2.09	2.08	2.15	5%	1%	**8%**
SDTHW	0.71	0.57	0.56	0.73	**0.61**	0.81	3%	6%	**43%**
MinHW	42.56	44.57	39.12	51.87	49.87	51.65	**22%**	12%	**32%**
MinTHW	1.19	1.23	1.20	1.17	**1.29**	1.18	-1%	5%	-2%

Note: * indicates a significant difference existing between experience groups in VTCF.

Mileage effects were very similar to that of driving experience, as drivers in all groups adapt their primary driving to the dual-tasking situation, it became clearer that drivers with the least mileage conducted more steering wheel adjustments, showed highest steering entropy and deviation in lane position, but also kept shorted headway, both in terms of distance and time.

6.5.2.4 Experience of and Attitude towards Visual IVISs

Performing visual tasks deteriorated driving performance for all groups. No effect of experience of using visual systems was found. For the attitude towards the manual IVISs, drivers who thought the systems are unsafe had the lowest speed ($F(2, 406) = 6.441$, $p = 0.002$) and the highest steering wheel adjustments (RR1.5, $F(2, 406) = 4.591$, $p = 0.001$; RR3, $F(2, 406) = 11.996$, $p < 0.001$) when performing visual tasks. For more details, see Table 6.19. Group #2 kept a significantly longer minimum headway, which explained why this group also reported a lower workload.

Table 6.19 Effect of attitude towards using manual systems on driving performance in VTCF

	BL			VT			VT Diff 1 & 3
Atti_Visu Group	#1	#2	#3	#1	#2	#3	
MNSP	83.97	85.74	84.08	84.03	84.65	**81.91**	**3%**
RR1.5	20.67	18.8	22.14	28.16	26.19	**30.93**	**-10%**
RR3	6.81	5.92	8.4	13.18	10.6	**17.02**	**-29%**

Note: * indicates a significant difference existing between groups when performing visual tasks.
* VT Diff 1 & 3 shows the difference between Group #1 and Group #3 when performing visual tasks, calculated by: (Value for #Group1 - Value for #Group 3)/ Value for #Group1.

6.5.3 Performing Auditory Tasks in Free-Driving

In the Free-Driving Scenario, no car following measurement was considered. In this and the next sections, the primary driving performance to be investigated includes the mean speed, SD speed, SD lane position, reversal rate per minute, and steering entropy.

6.5.3.1 Task Level

Performing auditory tasks in the Free-Driving Scenario did not impact the speed and lane keeping. There were noteworthy steering behaviour changes: RR1.5 and SE were higher when performing auditory tasks than in baseline (RR1.5, $F(3, 463) = 3.483$, $p = 0.016$; SE, $F(3, 463) = 2.627$, $p = 0.050$). The steering wheel turns increased from 28.29 times per minute in the baseline driving to 33.83 when performing the auditory tasks, as illustrated in Table 6.20. SE increases from 0.65 in baseline to 0.68 in task performing. The results suggest that the primary driving was affected very little by the auditory tasks in Free-Driving.

Table 6.20 Driving performance between baseline and auditory tasking in Free-Driving

Task Level	BL	Level 1	Level 2	Level 3	AT	Diff
MNSP	82.18	80.90	82.49	81.29	81.58	-1%
SDSP	7.29	8.30	7.43	8.21	7.97	9%
SDLP	0.18	0.17	0.17	0.17	0.17	-7%
RR1.5*	**28.29**	34.97	31.19	35.54	**33.83**	**20%**
RR3	9.32	10.88	10.27	12.30	11.12	19%
SE*	**0.65**	0.68	0.66	0.69	**0.68**	**5%**

Note: * indicates a significant difference existing baseline and performing visual tasks.

6.5.3.2 Gender Difference

Performing auditory tasks in Free-Driving Scenario had no significant effect on driving performance for male drivers, but affected female drivers remarkably. There was a decrease in speed (F (1, 148) = 8.119, p = 0.015), in crease in speed deviation (F (1, 148) = 8.679, p = 0.001), and increases in steering adjustments (RR1.5, F (1, 148) = 8.679, p = 0.004; RR3, F (1, 148) = 4.876, p = 0.029) and steering entropy (F (1, 148) = 7.569, p = 0.007).

When comparing the driving performance between two genders, female drivers kept a lower speed (F (1,381) = 15.124, p < 0.001), a higher SDSP (F (1, 381) = 5.357, p = 0.021), and more large steering wheel adjustments per minute (RR3, z = -2.380, p = 0.017), comparing to male drivers. Table 6.21 illustrates the detailed results of the driving performance measurements.

Table 6.21 Gender difference in driving performance in ATFD

Gender	Male			Female			Gender Diff in BLFD	Gender Diff in ATFD
	BL	AT	Diff	BL	AT	Diff		
MNSP*	82.09	82.27	0%	**82.30**	**80.00**	**-3%**	0%	**-3%**
SDSP*	8.10	7.62	-6%	**6.25**	**8.75**	**40%**	**-23%**	**15%**
SDLP	0.18	0.17	-7%	0.18	0.17	-7%	2%	1%
RR1.5	30.32	32.68	8%	**25.66**	**36.45**	**42%**	**-15%**	12%
RR3*	9.59	9.76	2%	**8.97**	**14.23**	**59%**	-6%	**46%**
SE	0.66	0.67	1%	**0.62**	**0.68**	**10%**	**-6%**	1%

Note: * indicates a significant difference existing between male and female in ATFD.

6.5.3.3 Experience/Age and Mileage

Performing auditory secondary tasks in Free-Driving did not affect driving performance on Experience Group #1 and #2. Experience Group #3 increased speed deviation (F (1, 124) = 4.696, p = 0.032), reversal rates (RR1.5: F (1, 124) = 12.896, p < 0.001; RR3: F (1, 124) = 6.037, p = 0.015) and steering entropy (F (1, 124) = 6.876, p = 0.010), and marginally reduced speed (F (1, 124) = 3.590, p = 0.060). However, when comparing the driving performance of the three experience groups, it was clear that Experience Group #1 showed a significantly lower driving performance. They kept the lowest speed ($F(2, 368) = 3.961$, $p = 0.020$), but showed highest speed deviation ($F(2, 368) = 2.848$, $p=0.050$) and higher large steering wheel adjustments ($F(2, 368) = 6.754$, $p = 0.001$). The detailed results are demonstrated in Table 6.22.

Table 6.22 Effect of experience on driving performance in ATFD

Experience Group	# 1		# 2		# 3		Diff BL vs. AT			AT Diff 1&3
	BL	AT	BL	AT	BL	AT	1	2	3	
MNSP*	81.17	**80.81**	81.36	82.62	83.88	81.76	0%	2%	-3%	**-1%**
SDSP*	7.76	**8.4**	7.58	7.26	6.57	8	8%	-4%	**22%**	**5%**
SDLP	0.19	0.17	0.18	0.16	0.18	0.17	-7%	-11%	-3%	0%
RR1.5	33.1	35.14	29.42	32.41	**22.09**	33.13	6%	10%	**50%**	6%
RR3*	12.68	**13.06**	8.5	8.36	**6.09**	10.84	3%	-2%	**78%**	**17%**
SE	0.66	0.68	0.67	0.68	**0.61**	0.66	3%	1%	**9%**	3%

Note: * indicates a significant difference existing between groups when performing auditory tasks.
* AT Diff 1 & 3 shows the difference between Group #1 and Group #3 when performing auditory tasks, calculated by: (Value for #Group1 - Value for #Group 3)/ Value for #Group1.

Similar results were found for age effect, where Age Group #3 adapted their driving behaviour for the auditory task. They slowed down ($F(1,145) = 4.918$, $p = 0.028$), had more steering adjustments (RR1.5: $F(1,145) = 13.306$, $p < 0.001$; RR3: $F(1,145) = 6.094$, $p = 0.015$) and higher SE ($F(1,145) = 7.755$, $p = 0.006$). Again, even after this compensation, the age effects still showed the same tendency when performing auditory tasks (SE, $Chi^2(2) = 7.754$, $p=0.021$), with younger drivers showing a higher SE, and the older ones having a lower SE.

Mileage showed similar effect as driving experience, where even Mileage Group #3 compensated the most, the least mileage group (#1) still showed the worst driving performance when performing auditory tasks.

6.5.3.4 Experience of and Attitude towards Auditory IVISs

Performing visual tasks deteriorated driving performance for all groups. No significant effect of experience of, or attitude towards, using auditory systems was found. The group who thought using auditory IVISs is not distracting (Group #1) kept a slightly higher speed (82.44kph versus 81.26 for Group #2, and 81.37 for Group #3), which resulted in higher steering entropy (0.7 versus 0.66 and 0.68 for Group #2 and 3 respectively), but the effects were not significant ($p>0.05$).

6.5.4 Performing Visual Tasks in Free-Driving

6.5.4.1 Task Level

It was found that performing the visual tasks while driving significantly decreased the speed (MNSP, $F(3, 445) = 9.463$, $p < 0.001$), increased speed deviation (SDSP, $Chi^2(3) = 39.871$, $p<0.001$), and changed steering behaviour (RR1.5, $F(3, 445) = 11.616$, $p < 0.001$; RR3, $Chi^2(3) = 63.983$, $p < 0.001$); SE, $Chi^2(3) = 34.823$, $p < 0.001$); a steady increase in SDLP across all levels was observed. This effect, however, was not statistically significant. The driving performance in all task levels as well as baseline are shown in Table 6.23. The Post Hoc Tests indicated that all differences only exist between baseline driving and the task-performing condition. No statistically significant change was found between 3 task difficulty levels. In the following analysis, the analysis is conducted between baseline and "visual tasking" condition.

Table 6.23 Driving performance between baseline and visual tasking in Free-Driving

Task Level	BL	Level 1	Level 2	Level 3	VT	Diff
MNSP**	82.18	78.94	78.17	78.55	78.54	**-4%**
SDSP**	7.29	10.36	10.91	10.53	10.61	**46%**
SDLP	0.18	0.19	0.19	0.2	0.19	6%
RR1.5**	28.29	38.18	39.07	38.26	38.52	**36%**
RR3**	9.32	18.59	20.25	18.95	19.29	**107%**
SE**	0.65	0.71	0.72	0.71	0.71	**9%**

Note: *denotes the p<0.05, ** denotes the p<0.01.

6.5.4.2 Gender Difference

When performing visual tasks, the primary driving performances for both male and female drivers changed dramatically compared to baseline driving. Both male and female drivers slowed down, increased speed deviation, made more steering wheel adjustments, and showed higher steering entropy (ANOVA, for all measurements, $p < 0.05$). The difference between the two genders while performing the visual tasks was not noticeable apart from female drivers keeping a slightly lower speed ($F(1, 352) = 5.541$, $p = 0.019$), and kept lower steering entropy ($F(1, 352) = 6.231$, $p = 0.013$), as shown in Table 6.24.

Table 6.24 Gender difference in driving performance in VTFD

Gender	Male			Female			Gender Diff in BLFD	Gender Diff in VTFD
	BL	VT	Diff	BL	VT	Diff		
MNSP*	**82.09**	**79.07**	**-4%**	**82.30**	**77.39**	**-6%**	0%	**-2%**
SDSP	**8.10**	**10.34**	**28%**	**6.25**	**11.19**	**79%**	**-23%**	8%
SDLP	0.18	0.19	8%	0.18	0.19	1%	2%	-4%
RR1.5	**30.32**	**38.10**	**26%**	**25.66**	**39.44**	**54%**	**-15%**	4%
RR3	**9.59**	**18.61**	**94%**	**8.97**	**20.78**	**132%**	-6%	12%
SE*	**0.66**	**0.72**	**8%**	**0.62**	**0.70**	**12%**	**-6%**	**-3%**

Note: * indicates a significant difference existing between groups when performing visual tasks.

6.5.4.3 Experience/Age and Mileage

When performing visual tasks in Free-Driving, Experience Groups showed slightly higher large steering wheel adjustments (RR3, F (2, 351) = 2.520, p = 0.082). Experience/Age Group #3 compensated their driving performance significantly. The compensation from the best performance groups, i.e. Experience/Age Group #3 almost eliminated the differences in baseline driving. For details, see Table 6.25.

Table 6.25 Effect of experience on driving performance in VTFD

	BL			VT			Diff BL vs. VT			VT Diff 1&3
Experience Group	1	2	3	1	2	3	1	2	3	
MNSP	81.17	81.36	83.88	78.17	78.69	78.98	**-4%**	-3%	**-6%**	-1%
SDSP	7.76	7.58	6.57	10.61	10.28	10.97	**37%**	36%	**67%**	-3%
SDLP	0.19	0.18	0.18	0.19	0.18	0.21	2%	1%	15%	-11%
RR1.5	33.1	29.42	**22.09**	38.31	37.8	39.64	**16%**	**28%**	**79%**	-3%
RR3*	12.68	8.5	**6.09**	**20.73**	17.57	18.83	**63%**	**107%**	**209%**	**9%**
SE	0.66	0.67	**0.61**	0.72	0.71	0.71	**8%**	**6%**	**17%**	1%

Note: * indicates a significant difference existing between groups when performing visual tasks.
* VT Diff 1 & 3 shows the difference between Group #1 and Group #3 when performing visual tasks, calculated by: (Value for #Group1 - Value for #Group 3)/ Value for #Group1.

While performing the visual tasks, Age Group #1 demonstrated higher steering entropy compared to Age Groups #2 and #3 (RR1.5, F (2, 351) = 5.706, p = 0.004; RR3: F (2, 351)=4.761, p = 0.009), and higher steering entropy (F (2, 351) = 7.428, p = 0.001). In performing visual tasks, each age group significantly reduced the speed, drives with higher speed deviation, and makes more RR1.5 and RR3 steering wheel turns. The increases of steering entropy were significant in Groups #2 and #3, but not in Group #13.

Mileage Group #1, who had the least annual mileage showed some risky behaviour comparing with other groups when performing visual tasks in Free-Driving. They showed a higher speed (F (2, 351) = 8.625, p < 0.001), higher lane deviation (F (2, 351) = 9.135, p < 0.001), companied by more large steering wheel adjustments (F (2, 351) = 5.667, p = 0.004).

6.5.4.4 Experience of and Attitude towards Visual IVISs

Performing visual tasks let to deterioration in driving performance for all groups. No significant effect for previous experience of using visual systems was found. However, in terms of *attitude*, the group who thought using manual IVISs is unsafe (Group #3) compensated more than the group who thought they are safe (Group #1) by a higher number of minor and large steering wheel adjustments (44.26 versus 37.35 for RR1.5, by 19%; and 25.53 versus 17.01 for RR3, by 50%, respectively, $p < 0.05$). The effort given to steering from Group #3 caused a lower SDLP (0.19 versus 0.2, by 5%), but this effect was not statistically significant.

6.6 The Relationship between Workload and Driving Performance

From Section 6.5, it can be seen that drivers' performance generally deteriorated along with an increasing workload, with fluctuation in some of their driving measurements. The region C in Meister' Theory (Meister, 1976), where performance stops decreasing and stays stable when the workload exceeds a certain level, was not observed in any of the four cases in this study. It is unlikely this was due to the workload not being high enough to reach the "plateau", because in the Post-Driving Interviews drivers stated that the tasks in Level 3 were already very demanding, and the workload involved in real driving is unlikely to be any higher, without potentially involving an accident. From the literature review (see Chapter 2) and the results from Section 6.5, the effects of auditory and visual tasks cannot be compared directly. For the auditory tasks, Figure 6.3 shows the relationship between workload and driving performance in the two Scenarios. This shows steering entropy increases with higher workload, suggesting unstable and deteriorated steering behaviour, although some fluctuations are observed in these relationships.

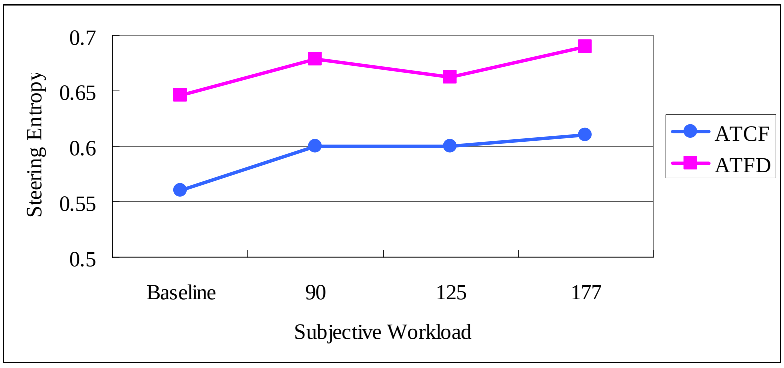


Figure 6.3 The relationship between workload and driving performance when performing auditory tasks

While performing the visual tasks, the deterioration in driving performance was more pronounced. Figure 6.4 shows the relationship between the subjective rated workload and the incremental increase in standard deviation of lane position compared to baseline. The solid line shows the observed value, and the dash line shows the anticipated value from the Spline Interpolation. The workload level of 100 states the equivalent of “turning on the radio”, and 200 would twice the workload of this task. With Car-Following, SDLP increases exponentially as the workload increased workload, suggesting an exponential decrease in driving performance. This is similar to Wewerinke and Hogema’s model (Wewerinke & Hogema, 2003), where SDLP increased exponentially with an increase in mental workload caused by higher speeds (Janssen, Brouwer & Huang, 2004; Wewerinke & Hogema, 2003), although this research was not specifically distraction-related. This finding is also consistent with previous research on the exponential relationship between the risk of accidents and deterioration factors such as alcohol/drugs consumption and fatigue levels (Brookhuis, 1998).

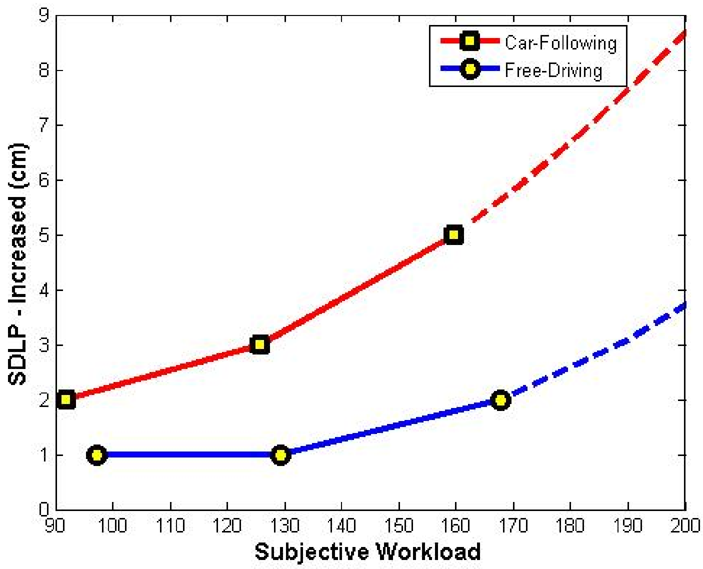


Figure 6.4 Relationship between workload and driving performance for visual tasks

6.7 Summary and Discussion

The main findings from this Chapter are:

- For both scenarios, drivers’ performance was impacted by both types of secondary tasks, but this was more pronounced in the visual ones.
- The auditory tasks caused an increase in minor steering wheel adjustments and higher steering entropy, indicating less predictable steering behaviour. While higher speed deviation was observed in both Scenarios, it was significant only in Car-Following. For the visual tasks, drivers also made more and abrupt steering, and showed higher deviations in both speed keeping and lane control, and they slowed down and increased their headway to compensate. As the workload increased in Car-Following, drivers’ performance was found to deteriorate exponentially while performing the visual tasks.

- Comparing genders, the auditory tasks affected female drivers, but seemingly little for male drivers. For the visual tasks, female drivers were also more affected. However, they compensate more than male drivers for both the auditory and visual workload, and showed more stable vehicle control, suggesting a more conservative coping strategy.
- Drivers with less than 10 years' experience showed worst driving performance both in baseline driving and while performing the secondary tasks, and this effect was more pronounced for the visual tasks. Drivers with the most experience were shown to compensate more for the extra workload, due to a higher developed capacity and maybe also increased awareness. The driving performance of the least experienced Group deteriorated even further with higher workloads, but they did not compensate to the same degree as the more experienced ones.
- Experience of using IVISs had no significant effect on driving performance. However, drivers who thought IVISs were either distracting (auditory IVIS) or unsafe (manual IVIS) showed more conservative behaviour while performing the secondary tasks.
- The effects between the two Driving Scenarios were consistent, although the results from Car-Following were more consistent and conclusive.

Details of these findings are discussed and explained below.

6.7.1 The Effects of Secondary Tasks on Driving Performance

Driving behaviour was generally affected in both Scenarios while performing the secondary tasks, especially the visual ones. In the Car-Following Scenario, there was an increase in speed *deviation* for all drivers while performing the auditory tasks of 19% (average of 2.5kph versus baseline of 2.1), and a general increase in minor steering wheel adjustments of 20% (average of 25 times per minute versus baseline of 21), which is consistent with previous on-road findings (Engström, Johansson & Östlund, 2005; Östlund et al., 2004). Less predictable steering behaviour was also found (a higher steering entropy of 9%), which suggest drivers' behaviour were impacted by the auditory tasks, but this may not be significant in every case. More significant driving performance changes were found in the visual tasks, with an increase in lane deviation by 18% (average of 18cm versus baseline of 15), increases in both minor and large steering wheel adjustments (by 36% and 85% respectively), and more unpredictable steering behaviour (Entropy increased by 13%). Visual tasks also resulted in significantly increased headways, with the mean distance headway increased by 24% (59.6m compared to baseline of 48.1m), mean time headway increased by 5.2% (2.11s compared to 2.00s) and both the *minimum* distance and time headways increased also, which suggests that the performance of all drivers was significantly affected by performing these tasks.

Similar performance changes were found for drivers in the Free-Driving Scenario. There were some effects while performing the auditory tasks, with a general increase in minor steering adjustments of 20% (average of 33.8 times per minute versus baseline 28.3), and higher steering unpredictability (Entropy increased by 5%). The visual tasks had a profound effect on all drivers, with a higher speed deviation (by 46%, from 7.29 to 10.61kph on average) *and* a decrease in average speed (from 82.18 to 78.54kph), which suggest more caution whilst driving, yet still less stable *speed* control. There were also significant increases in minor and large steering wheel adjustments (by 36% and 107% respectively) and higher steering unpredictability (entropy increased by 9%), which all suggest a deterioration in driving performance.

For the visual tasks, where the conflicts of visual and manual resource were very high, drivers chose either to increase their headway (while Car-Following) or slowed down where possible (while Free-Driving) to compensate. The effects of visual task on headway and speed have been observed in other research (Engström, Johansson & Östlund, 2005; Östlund et al., 2004), and were noted (Burns et al., 2002) as strategies applied by drivers to cope with the competing demands of dual-tasking. The increased steering adjustments for both Scenarios also showed the extra effort for drivers to keep lane control under the higher workload caused by dual-tasking, especially in the frequency of large steering wheel turns, and less predictability in steering control, which suggest drivers were trying to adjust to the higher task demand. Similar findings were reported by Burnett and Joyner (Burnett & Joyner, 1997), where higher steering wheel variability was found for drivers who looked at a route guidance system while approaching dual-carriageway exits, which was explained as compensatory adjustments by drivers to correct any path deviation caused by looking away from road ahead. However in this study, even with extra effort applied in steering, a higher lane deviation in Car-Following was still observed for the visual tasks, indicating deterioration in lane control. A similar compensation effect was observed for speed control in Free-Driving, with drivers generally slowing down, and a greater speed deviation again observed for the visual tasks. The workload involved in using a complex IVIS system is likely to include a combination of mental and visual elements. Given that visual tasks had a more pronounced effect on driving performance, due to the high conflicts of visual and manual resource that are involved, it is likely the effect of a combined workload will be dominated by the visual component. Therefore, looking for ways to detect mental distraction in this context is important, because according to drivers' Subjective Rating, the workload from both elements can be equally demanding (see Chapter 5).

These effects were principally observed in dual-tasking compared to baseline driving, but the variation between task difficulty levels was *not* statistically significant, which were also observed in previous distraction research (Engström, Johansson & Östlund, 2005; Östlund et al., 2004). This suggests the workload increase *between* difficulty levels was less material to that of *adding* a secondary task whilst driving, and could be explained by drivers adopting *different* coping strategies across the three task levels, which would support Vicente et al. (Vicente, Thornton & Moray, 1987)'s theory that drivers may invest different effort through the difficulty levels as compensation. In addition, at the higher difficulty levels, drivers may also have prioritised their resources so that the primary task of driving took precedence over all other tasks, and this accounts for the deterioration in *secondary* task performance as the difficult level increased, which is referred to previously (Section 5.5), and supports Wicken's theory (Wickens, 1984a) in the literature.

6.7.2 Effects of Driver Characteristics on Driving Performance

6.7.2.1 Gender Effects

A significant difference in driving performance was observed between male and female drivers while performing the secondary tasks in both Scenarios. For the auditory tasks, female drivers exhibited greater changes in driving performance and behaviour compared to baseline, while male drivers were hardly affected. For the visual tasks, both males and females were similarly affected, and this resulted in driving performance deterioration in both cases.

In Car-Following, female drivers increased their minor and large steering wheel adjustments while performing the auditory tasks (by an average of 37% and 46% respectively compared to baseline), and their steering entropy also increased by 14%, while there were no significant steering changes in male drivers. In turn, female drivers showed better lane control, with 10% lower deviation on average compared to males (of 13.6 and 15.1cm respectively), which may reflect their increased effort in steering. Female drivers also increased their distance headways (with both the mean and minimum increased by 20% compared to baseline), but this effect was not observed in male drivers. These driving behaviour changes suggest female drivers apply compensation strategies to account for the extra task demand from auditory tasks, which resulted in more stable lane control. Male drivers on the other hand, did not exhibit behaviour changes apart from an increase in speed deviation (by 14% compared to baseline, $p = 0.032$), which suggest they either did not apply similar compensation strategies or they were not impacted by the demand from auditory tasks to the same extent or felt the need to compensate for them, although they showed higher deviation in longitudinal control (22% higher standard deviation in headway than females).

The driving performance of both genders deteriorated for the visual tasks while Car-Following, with both male and female drivers increasing their headways to compensate. For female drivers, the minor and large steering wheel adjustments and steering entropy all increased (by 43%, 131% and 16% compared to baseline respectively), as did those for male drivers but to a less extent (by 32%, 65% and 11% respectively). However, female drivers managed to maintain more *stable* headways compared to males (22% less headway deviation) *and* a lower lane deviation (8% less), which along with the increase in steering adjustments, suggest they compensate for visual tasks even more, and this resulted in more stable lateral *and* longitudinal control, although their performance also deteriorated generally compared to baseline.

Similar gender effects were found for the Free-Driving Scenario. The driving behaviour of female drivers was more affected by performing the auditory tasks than males, again suggesting they compensated for these tasks, including slowing down (80kph compared to baseline of 82.3kph, $p = 0.015$) and increasing minor and large steering wheel adjustments (by 42% and 59% compared to baseline respectively), while male drivers were not affected. A similar deterioration in speed *control* was found when drivers used a hand-held phone (Burns et al., 2002), which resulted in a neglect to monitor speed and therefore were unable to maintain their normal driving speeds. A higher speed *deviation* (40% more than baseline) was also found for female drivers in the auditory tasks, while males were not generally affected, which suggest speed control is generally *harder* for them to compensate for, due to a reduction of spatial references in Free-Driving compared to e.g. Car-Following, which is consistent with Petrevu's research (Putrevu, 2001) that showed biological differences in information processing between genders.

Driving performance deteriorated for both genders for the visual tasks in Free-Driving, with both male and female drivers reducing their speed to compensate. Again, for female drivers, the minor and large steering wheel adjustments and steering entropy all increased compared to baseline (by 54%, 132% and 12% respectively), which were more than the increase for male drivers (26%, 94% and 8% respectively). In this case, female drivers maintained more stable *steering* (3% less in Steering Entropy), which along with higher steering adjustments and lower average speed (73.39kph compared to 79.04kph for male drivers), suggest they again compensate more than the males for the visual tasks, although both genders deteriorated compared to baseline.

In conclusion, female drivers appear to compensate more than males while performing both types of secondary tasks and in both Scenarios, by increasing their steering adjustments, and either slowing down (Free-Driving) or increasing their headway (Car-Following), which concurs with Dorn's theory that females generally want to drive more attentively than males (Dorn et al., 1991). This gender difference was more pronounced in the auditory tasks, where there was little behaviour change in male drivers. When using a mobile phone, Faulks et al suggest (Faulks et al., 2009) females appeared to be more distracted in driving than males. However in this research, female drivers showed less deterioration in lateral control and more stable headway control (in Car-Following) or more stable steering control (in Free-Driving), which suggest they applied a more *conservative* coping strategy in order to gain higher stability. Male drivers on the other hand showed better speed control in Free-Driving, where they were instructed to keep a constant speed.

6.7.2.2 Effect of Driving Experience/Age and Mileage

Experience and annual mileage both had a significant effect on driving performance. Age had a similar effect to experience/mileage, but was less significant. These effects were apparent both in baseline driving and when performing secondary tasks, although more pronounced for the visual tasks.

In *normal* or baseline driving, the most experienced Group (#3, with 20+ years driving experience) clearly performed the best. For example, in Free-Driving, they had lower steering adjustments and Entropy compared to Group #2 (10-20 years driving experience) and #1 (less than 10 years experience), with e.g. an RR1.5 of 22.09 compared to 29.42 and 33.1 respectively. For some measures, the performance between Groups #3 and #2 were broadly similar, but the performance for the least experienced Group (#1) was significantly *worse* in all cases - for example, their lane deviation in Car-Following was much worse, with an average of 18cm compared to 14 and 13 for Groups #2 and 3. They also had higher minor and large steering wheel adjustments (e.g. 45% and 102% more than Group #3 in Car-Following, and 33% and 54% more in Free-Driving), and higher Entropy (e.g. 11% more in Car-following).

The differences *between* the Groups continued when comparing their relative driving performance in the auditory and visual tasks for both Scenarios, i.e. the performance of the *more experienced* Groups, #2 and #3, were broadly comparable, with Group #3 being better in some cases. For example, although steering adjustments were slightly higher than Group #2, their Steering Entropy was generally lower for both Car-Following and Free-Driving, but none of these differences were significant. The least experienced Group #1, again performed worst in all cases. For example, their lane deviation in Car-Following was 10% worse than the most experienced Group while performing the auditory tasks, and 15% worse for the visual tasks;

with more *large* steering adjustments in Free-Driving (17% more for auditory tasks, and 9% for visual) and higher Entropy than Group #3.

Comparing the Groups' performances in the secondary tasks *against* their own baselines, all drivers compensated for the higher extra workload of *visual* tasks in both Scenarios, i.e. they either *increased* their mean headways (in Car-Following) or *reduced* their mean speed (in Free-Driving), but this still resulted in higher Steering Entropy, and a lane keeping deterioration (in Car-Following) or higher speed deviation (in Free-Driving) in all cases. The compensation for visual tasks was exhibited the most by Group #3, which suggests they were more aware of the extra demand and potential risks imposed by the visual tasks. This is also reflected in their performance of auditory tasks, where Group #3 increased either their mean headway (Car-Following) or reduced their average speed (Free-Driving) relative to their baseline to account for the extra auditory task demand.

Even with these compensation actions, the performance of Group #3 was still better than Group #1 (where there was little such behaviour change), and broadly comparable with Group #2 (where the change was minor). This suggests more experienced drivers have a higher *capacity* to compensate for the extra workload involved in dual-tasking. According to the *skill-rule-knowledge* theory (Rasmussen, 1987) or *control-manoeuvring-strategy* in driving (Michon, 1985) - see Chapter 2, more experienced drivers are able to *automate* the vehicle control (required at the *skill or control level*) so that this releases extra resources, which helps to improve their decision making (required at the *rule or manoeuvring level*), and this increases their capacity for performing tasks in complicated situations, such as dual-tasking when driving. For example, when performing the visual tasks in Car-Following, Group #3 increased their minor and large steering adjustments by 54% and 135%, and their mean headway by 33%. In contrast, Group #1 increased their steering adjustments by only 20% and 55% respectively, and their headway by 26%; and while lane deviation deteriorated for both Groups, the value for Group #3 when performing the visual tasks was still less than the *baseline* for Group #1.

In contrast, the *least* experienced drivers (Group #1) may not be so aware of the *extra* workload and risks imposed by the secondary tasks, and therefore did not adapt their driving behaviour to the same degree. This is reflected in their Subjective Ratings (see Chapter 5), where they rated the increased workload from the secondary tasks as being significantly *lower* than the other two Groups. These drivers were also less aware of the *scale* of the workload, which is especially noticeable when performing the auditory tasks. For example, in Free-Driving, they significantly reduced their average speed when performing the visual tasks, but not so for auditory - whereas the most experienced drivers did so for both tasks. This resulted in a worse performance for *both* the auditory tasks (see Chapter 5) *and* driving (e.g. a significantly *higher* speed deviation than the other Groups). In fact, the least experienced Group showed the *worst* performance in both the primary and secondary tasks *overall*, which suggests they could *not* have prioritised one task over another - a possible coping strategy suggested by Wickens (Wickens, 1984a), see Chapter 5 - as may have been the case for Groups #2 and 3. While it is possible they were not so engaged in performing the primary and/or secondary tasks, it is more likely they were less aware of their impacts, and therefore had a lower capacity to deal with them.

Contrary to expectation, while the behaviour change or coping strategy applied by the most experienced drivers (Group #3) generally resulted in better *driving* performance than Group #1, it was *not* necessarily better than Group #2, who also showed a slightly better *secondary* task performance in some cases (see Chapter 5), which suggests that Group #3 drivers may not just prioritise the primary task of driving to cope with the higher workload, and may therefore apply a different coping strategy to Group #2. As both the primary and secondary task performance for Groups #2 and 3 were broadly similar, it is likely that the *benefit* of driving experience does not increase significantly beyond a certain threshold, except in normal baseline driving - at least according to the driving performance data alone. However, it is still possible that, with even more experience, drivers paid more attention to their surroundings, and therefore achieved the best possible balance between safety and secondary task performance while driving (see Chapter 7). This would extend existing driving behaviour models (Michon, 1985; Rasmussen, 1987) to suggest that, with significantly increased experience, behaviour at the manoeuvring level can escalate beyond increasing the capacity for performing tasks in complication situations, to adopting a more effective strategy in driving.

Age showed a similar but less significant effect on performance than driving experience. For example, older drivers (Age Groups #2 and 3) compensated for secondary tasks the most, while younger drivers (Group #1) showed greater lane deviation in Car-Following (16cm, compared to 15 and 14 for Groups #2 and 3) and higher steering entropy in Free-Driving (0.69 compared to 0.65 and 0.60), but the difference was not so pronounced.

The findings for mileage were similar to those of driving experience, although the two characteristic factors were not significantly correlated. For example, while performing auditory tasks in Car-Following, no significant behaviour change was found for the least Annual Mileage Group (#1), while those with higher mileage (Group #2 and #3) exhibited greater compensation. For example, Group #3 increased their minor and large steering wheel adjustments the most (by 34% and 27%).

6.7.2.3 Effect of Experience/Attitude towards Using IVISs

Previous experience of using either auditory or manual IVISs had no significant effect on driving performance, even though it improved the accuracy of secondary tasks (see Chapter 5). However, drivers who thought using IVISs whilst driving as either distracting (auditory IVIS) or unsafe (manual IVIS) showed more conservative behaviour while performing the secondary tasks than those who did not. In Car-Following for example, Atti-Audi Group #3 drivers (who considered auditory IVISs distracting) showed a longer headway while performing the auditory tasks than Group #1 (who thought they are not distracting), i.e. 53.7m versus 44.6m or 20% longer, without any other deterioration in driving performance. Similarly, Atti-Visu Group #3 drivers (who thought manual IVISs unsafe) were more conservative while performing the visual tasks - they kept a lower speed (81.9kph versus 84 for Group #1), and made more effort with highest minor and large steering wheel adjustments (10% and 29% higher than Group #1 respectively), without sacrificing lane control. Similar effects were found for the different driver attitudes in Free-Driving, where Atti-Audi Group #3 drivers kept a lower speed (82.44kph versus 91.37 for Group #1) in auditory tasking, with lower steering entropy (0.68 versus 0.7 for Group #1) as a result; and Atti-Visu Group #3 drivers made more minor and large steering wheel adjustments (by 19% and 50% higher than Group #1 respectively) to keep better lane control (by 5% than Group #1) while visual tasking. This suggests that drivers from Group #3 were more cautious of the extra workload caused by in-vehicle tasks, and slowed down or make more effort in steering to deal with dual-tasking, and to release mental and physical resources (e.g. ability to conduct more steering wheel adjustments) for better vehicle control and more stable lane position; and this resulted in a lower Subjective workload Rating when performing the secondary tasks, and a fast reaction time for the visual ones (see Chapter 5). This also explained why drivers who thought IVISs were safe appeared to experience a higher subjective workload, compared to those who were against, because they had a more relaxed attitude towards dual-tasking and did not adopt the conserving approach.

Chapter 7 Impact of Secondary Tasks on Drivers' Visual Behaviour

7.1 Introduction

This Chapter presents the findings of the impact of in-vehicle secondary tasks on drivers' eye movement. In this Chapter, the eye movement of drivers in the Car-Following Scenario only was analysed and discussed, as previous chapters (see Chapter 5 and 6) have shown that the effects of secondary tasks on driving performance were broadly consistent across the two driving Scenarios; and the measurements were more sensitive and results more significant in Car-Following.

Similar to Chapter 6, a comparison of drivers' eye movement was made between the three task difficulty levels and baseline driving for both the auditory and visual tasks in Car-Following (or ATCF and VTCF respectively). Again, the effect of different driver characteristics, i.e. gender, age, driving experience, annual mileage, experience and attitude towards using IVISs, on drivers' visual behaviour when performing these secondary tasks were investigated, along with different coping strategies and drivers' attention allocation, which provides further information on the relationship between the increasing workload and drivers' performance changes across the different characteristics groups.

7.2 Eye Movement Data Preparation

As described in Chapter 4, eye movement or visual behaviour data were collected using the FaceLAB™ System with a frequency of 60 Hz. The System collects drivers' eye movement data in four categories, which are indexed by *frame number* for each test run. The categories are:

- time stream data;
- head position and rotation angle measurements;
- eye position and rotation, and *blink* and *saccade* (or endogenous eye movement measurements); and
- world model data, which contain the *gaze objects* or gazes areas of interest.

The eye movement data were re-organised using programs written by the author into a structured database with a sample rate of 60Hz, and indexed by time (see Chapter 4). As the data collected from the FaceLAB™ System contained a certain degree of noise, the original data were filtered for each subject using criteria based on human eye movement behaviour, which is described in Section 7.2.1. For more details on the nature of human eye movement, including the theoretical basis for the criteria used, see Chapter 2 (Literature review). After filtering, each eye movement parameter was extracted and calculated for each *task-episode entry* (as discussed in Chapter 4), and analysed.

7.2.1 Criteria of Eye Movement for Data Filtering

Traditionally, the analysis of human beings' eye movement focussed primarily on the separation of their *fixations* with *saccades* (Crundall et al., 2004), which are the two key aspects of visual information extraction. However, the literature provides different opinions on how fixations and saccades are defined and differentiated (Nakayama et al., 1999; Putrevu, 2001). In essence, fixation concerns the focus of gaze on objects for information extraction, during which the eyes are relatively stationary, while saccade is the eye movement between fixations, which lead the gaze from one object to another. Saccades are therefore rapid eye movements, with peak gaze velocities typically ranging from 300°/s to 700°/s (Putrevu, 2001), and at such speed, the brain is no longer able to perceive the visual information intake. Fixations, on the other hand, typically involve pauses over objects where the eyes can absorb information, and only gazes longer than 140 ms are related to focal processing (Summala, 1981). *Smooth pursuit* is a sub-category of fixation commonly found in driving, i.e. fixation to track slow moving objects, with a gaze velocity that ranges from 80°/s to 160°/s (Putrevu, 2001). See Chapter 2 for further descriptions of eye movement, including physical characteristics, regions of visual field, and how they related to information processing and behaviour organising.

The separation and selection of fixation and saccades used in this study were therefore based on temporal criteria, with the former being longer than 140 ms in duration and with a gaze velocity lower than 300°/s; while the latter ranged from 10 to 99 ms. Unlike eye movement in reading, the fixations involved in driving can be much longer in duration, with smooth pursuits being included as one of the major types of fixation (Larsson, 2002; Victor, Harbluk & Engstrom, 2005). Therefore, no upper duration limit for fixation was defined this study, and movements outside the ranges mentioned were ignored, in common with other studies (Di Stasi et al., 2009; Rayner, 1998).

7.2.2 Validation of Gaze Objects of Interest

In addition to gaze angle filtering for saccade and fixation, the gaze objects of interest used by FaceLAB^{TM} were validated for each driver. This process is described below.

7.2.2.1 Validation Method and Reasoning

The sizes of gaze objects of interest and their locations were defined in the *world model* within the FaceLAB^{TM} System for each individual driver, according to their specific head location, face features and gaze movements (see Section 3.4.2). FaceLAB^{TM} then provides a frame-by-frame estimation of the *gaze objects* or where a driver is looking at in each frame, using their eye positions, gaze angles, head angles and positions, relative to the locations of the pre-defined objects. These *gaze objects* readings from the System were then mapped and further validated using the *Naive Bayes Classifier* (see below). The gaze objects validated are:

- Right Mirror, identified as Object Number 2;
- Rear Mirror, identified as Object 5;
- Tachometer, identified as Object6;
- Speedometer, identified as Object 7;
- Touch Screen, identified as Object 8, and
- to the Front, identified as Object 9.

These gaze objects of interest were validated for each driver as follows:

1. The equipment was calibrated for each subject before driving started, by asking them to look at each object in the World Model, and the gaze objects identified by the System, and the accuracy of gaze inspected visually on the FaceLAB ^{TM} display, which showed images of the objects locations and real-time gaze angles, as well as the interaction between the two.
2. Instructions were then given to drivers to briefly look towards the gaze objects at the start of each driving test run - for example, "quickly check the right mirror", and the interactions between the driver's gaze and the gaze object visually inspected again.
3. During the in-vehicle tasks, the time when certain actions were taken by the driver, e.g. when looking at touch screen, was also recorded manually by the experimenter, and used to compare against the data derived from the FaceLab^{TM} System to ensure the correct gaze object was being recorded.
4. A Naive Bayes Classifier was then used to map and determine whether FaceLab^{TM}'s estimations of drivers' gaze objects were consistent with the actual position of the objects and the driver's eye-head motions over time. The Naive Bayes Classifier is a *probabilistic classifier* (Duda & Hart, 1973) used in machine learning, based on **Bayes' Theorem** with assumptions of independence. The classifier can be trained to establish a probability algorithm efficiently using a supervised *learning* method, which can then be used to *predict* other instances (See Section 7.2.2.2 below for more details). The

gaze objects from FaceLAB^{TM} were confirmed as reliable when the following criteria are met:

- the overall *error* of the original gaze objects read from FaceLAB^{TM}, as compared to the predicted results from the Classifier, are at a low level, i.e. the FaceLAB^{TM} gaze objects and those predicted by the Classifier should be consistent, with only small errors occurring at the boundaries;
- the gaze objects from the System and the Classifier are both consistent with the head and eye movements of the driver relative to the objects, i.e. the movement angles in FaceLAB^{TM} should reflect real motions driving; and
- this is reflected in the data through a noticeable increase in time spent looking at the touch screen when drivers performing the visual task.

Principle Factor Analysis was conducted for *dimensionality reduction* or to select important variables for use in the Naive Bayes Classifier, as the potential input variables were *multi-dimensional* (e.g. they include eye as well as head movements, and the relative locations of objects) and could be closely correlated. From a preliminary analysis, it was found that the head data was sometimes more stable for gaze objects than the eye motion data. Therefore both the head and eye movement data from FaceLAB^{TM} were included in the validation process. Potential variable groups tested for input were:

1. 3 head position variables (in x, y, z axis);
2. 3 head rotation angle variables;
3. 3 head position variables + 3 head rotation angle variables;
4. the pitch and yaw angles of eyes (or vertical and horizontal angles); and
5. pitch and yaw angles of eyes + 3 head position variables + 3 head rotation angle variables.

The results from the Principle Factor Analysis showed that variable group #2 (head rotation angles) proved to be the most effective in explaining the variance in the head object data, while group number 4 (pitch and yaw angels of eyes) accounted for the variance in eye gaze object data, and they were therefore selected as the input variables for the Classifier used to train the prediction algorithm for gaze objects.

7.2.2.2 The Naive Bayes Classifier

The Naive Bayes Classifier uses a *supervised learning* algorithm, as originally proposed in *pattern recognition* (Duda & Hart, 1973). The supervised learning method infers a function from training data (i.e. prior knowledge) which contains an input (e.g. eye or head variables in this study) and an output (e.g. the gaze objects), and extracts and stores from it a probabilistic *summary* which contains the conditional probability of each attribute value for each *class* (e.g. the categories of gaze objects), as well as the probability of each class, by analysing the known *instances* (i.e. cases from training data). Through the training process, each time a new *instance* is fed into the Classifier, it updates the summary and the boundaries of the instance space. A probability *algorithm* is then established after the training process, which is used for the prediction process. When a *test instance* is given to the Classifier for prediction, it uses an evaluation function that contains the algorithm to *rank* the test instance, and assign this instance to the class which has the highest probability.

The Naive Bayes Classifier requires the attributes of the instances to be independent, although this assumption is often violated in practice. However, despite this, others suggest (Zhang, 2004) that the Naive Bayes Classifier still performs well.

The Naive Bayes Classifier contains a conditional probability model, i.e.:

$$p(C|F_1,...,F_n) \ , \qquad (7.1)$$

where C is a dependent class variable, conditional on several features of the variables F_1 to Fn.

According to Bayes' Theorem:

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}. \qquad (7.2)$$

Therefore the model in (7.1) can be described as:

$$p(C|F_1,...,F_n) = \frac{p(C)\,p(F_1,...,F_n|C)}{p(F_1,...,F_n)}. \qquad (7.3)$$

In practice, only the numerator in this formula is of interest, since when the probability of the features F_i are known, the denominator is effectively constant and not depend on C, and the numerator therefore is equivalent to the joint probability model, i.e.:

$$p(C, F_1,...,F_n) \qquad (7.4)$$

which can be rewritten as follows, using the definition of conditional probability:

$$
\begin{aligned}
& p(C, F_1, ..., F_n) \\
& = P(C)\, p(F_1, ..., F_n \,|\, C) \\
& = p(C)\, p(F_1 \,|\, C)\, p(F_2, ..., F_n \,|\, C, F_1) \\
& = p(C)\, p(F_1 \,|\, C)\, p(F_2 \,|\, C, F_1)\, p(F_3, ..., F_n \,|\, C, F_1, F_2) \\
& = p(C)\, p(F_1 \,|\, C)\, p(F_2 \,|\, C, F_1)\, p(F_3 \,|\, C, F_1, F_2)\, p(F_4, ..., F_n \,|\, C, F_1, F_2, F_3) \\
& = p(C)\, p(F_1 \,|\, C)\, p(F_2 \,|\, C, F_1)\, p(F_3 \,|\, C, F_1, F_2) ... p(F_n \,|\, C, F_1, F_2, F_3, ..., F_{n-1}).
\end{aligned}
\quad (7.5)
$$

Based on the *naive* condition (i.e. independence assumption) which assume that each feature F_i is independent of every other feature F_j , when $i \neq j$, the model can be further developed to:

$$p(F_i \,|\, C, F_j) = p(F_i \,|\, C) \quad (7.6)$$

I.e. for $i \neq j$, the joint model can be expressed as:

$$
\begin{aligned}
& p(C, F_1, ..., F_n) = p(C)\, p(F_1 \,|\, C)\, p(F_2 \,|\, C)\, p(F_3 \,|\, C) ... \\
& = p(C) \prod_{i=1}^{n} p(F_i \,|\, C)
\end{aligned}
\quad (7.7)
$$

This means that under the above independence assumptions, the conditional distribution over the class variable *C* can be expressed as:

$$p(C, F_1, ..., F_n) = \frac{1}{Z} p(C) \prod_{i=1}^{n} p(F_i \,|\, C) \quad (7.8)$$

where *Z* is only dependent on $F_1, ..., F_n$, i.e., a constant if the values of the feature variables are given.

The Naive Bayes Classifier combines this model with a *decision rule*, and one common rule used is to choose a hypothesis that has the highest probability, known as the *maximum posteriori* or *MAP* decision rule. The corresponding classifier is therefore the function defined as follows:

$$\text{classify}(f_1, ..., f_n) = \arg\max_c \, p(C = c) \prod_{i=1}^{n} p(F_i = f_i \,|\, C = c). \quad (7.9)$$

7.2.2.3 Validation Results from Classifier

Because of significant individual differences between drivers, a Classifier model was established for each driver. Half of the visual behaviour data collected in each test run (i.e. from each driver) was used to train the Naive Bayes Classifier model, and the other half used as *test instances* for the prediction, i.e. to verify the FaceLAB™ gaze object prediction. The partition of the two half-datasets was based on taking every other data entry from the whole dataset for one half, and leaving the rest for the other half. The accuracy of prediction was then evaluated using each combination of input variables (as described in 7.2.2.1) by visual inspection and checking the percentage error.

After an analysis of the various tests, the gaze objects obtained through FaceLAB^{TM} showed similar results to those predicted from the Naive Bayes Classifying, where the head rotation angles were shown to be able to predict the head objects with an average prediction error of 0.8%, while the two eye gaze angles were able to predict the gaze objects with a prediction error of 1.52%. Therefore, the validation criteria discussed in Section 7.2.2.1 were verified, and the gaze objects data collected by FaceLAB^{TM} deemed valid and reliable.

Figure 7.1 shows the head gaze object based on the three head rotation angle parameters for one subject, and Figure 7.2 shows the gaze object based on the two gaze angle parameters for the same subject. The ellipsoids and ellipses in these figures are the error ellipses of the co-variances based on a 95% confidence interval. These figures visualised the conclusion of these parameters being good input variables for predicting objects of interest. The estimation of gaze objects from the FaceLAB^{TM} is consistent with the position of the objects and eye-head motions, and is therefore deemed reliable.

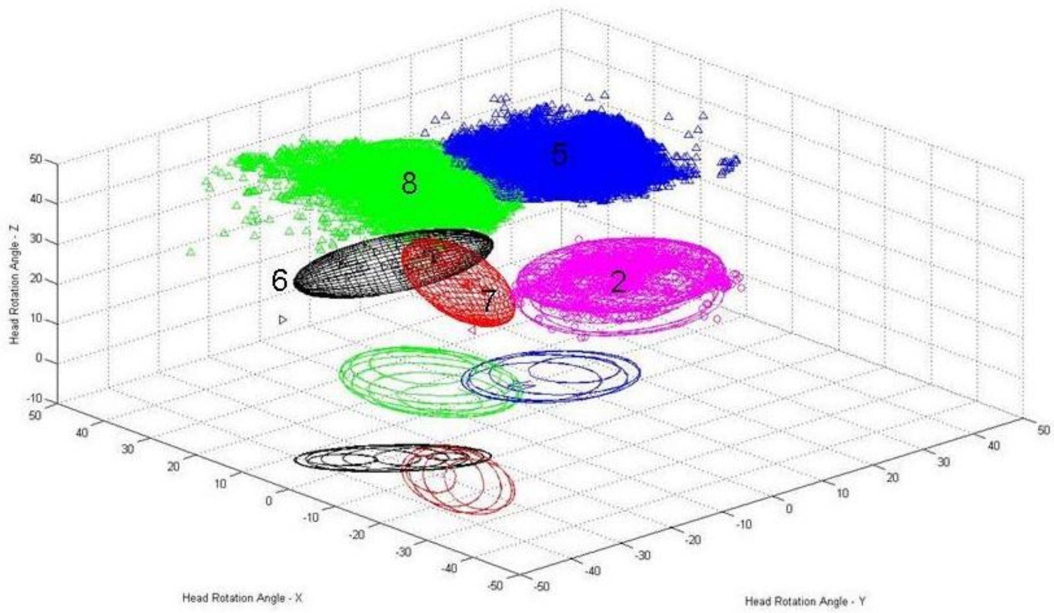


Figure 7.1 Error ellipses for the areas of interest in head-coordination

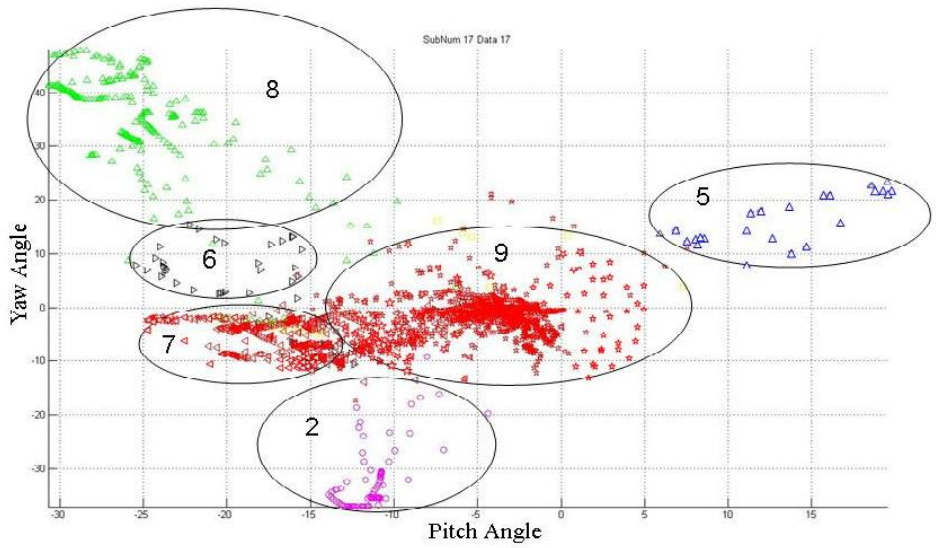


Figure 7.2 Error ellipses for the areas of interest in eye-coordination

Note: in the error ellipses, a 95% confidence interval was applied

7.3 Parameters Description

The following parameters were used to describe drivers' visual behaviour:

7.3.1 Percentage of Gaze on Areas of Interest

The percentage of time drivers spent gazing at specific "objects" or *areas of interest* was calculated in order to understand how their mental and visual resources are managed, and how their attention shifted between different areas during dual-tasking versus baseline driving. The different areas of interest are:

1. *Front* (or windscreen): gaze on the road ahead;
2. *Mirrors*: includes gaze on the right mirror and the rear-view mirror. The left mirror was not included because its location is far away from the "Stereo-head", the two cameras used by FaceLAB™ (item number 4 in Figure 3.10), which therefore could not be tracked accurately.
3. *Panel*: gaze on the dashboard or in-vehicle panel, including the tachometer and speedometer.
4. *Touch Screen*: used by drivers to perform the visual tasks, and
5. *"Nothing"*: representing the time spent looking at other areas of interest, which are peripheral to those listed above.

Drivers tend to spend most time looking at the front (Hughes & Cole, 1986). The Front in this study was defined by the size and location of Windscreen relative to each individual driver, but not based on fixed gaze angles as it was in the Victor's field study, where Front area was defined as a circle of 16°diameter, centered around the road centre point (Victor, Harbluk & Engstrom, 2005). After the experiment, the gaze range for Front was calculated, and it was found 99% of the gazes in Front were within the range of 12°diameter centered around the central vision. Drivers also look at the mirrors (Pastor et al., 2006) to obtain driving-related information, to adapt their driving to the traffic conditions. In addition, the time spent looking at *nothing* reflects the frequency and/or duration of gaze transition *between* the different areas of interest, which reflects the *efficiency* of visual resource deployment, i.e. how much time is spent switching between areas (versus focussing on one), and therefore the visual *resource cost* (Rogers & Monsell, 1995).

The shifting of attention between different areas of interest could also reveal whether drivers use different strategies to cope with the extra demands of in-vehicle tasks. For example, a *higher* percentage of time on mirrors can reflect a compensation strategy being used by drivers when the workload is high, suggesting a desire to obtain more driving information, and therefore (Brookhuis, De Vries & De Waard, 1991) increased alertness to the traffic conditions.

7.3.2 Standard Deviation of Gaze

Standard deviation of gaze has been shown to be sensitive to increasing mental workload (Victor, Harbluk & Engstrom, 2005). With similar concept, spatial gaze variability was used by Recarte and Nunes (Recarte & Nunes, 2003), which represents an area of a rectangle of the Standard Deviation of vertical gaze coordinates multiplied by the Standard Deviation of horizontal gaze coordinates.

In this study, Standard Deviation for Gaze (named as SD Gaze) is defined as the standard deviation of the combined vertical and horizontal gaze angles. Pythagoras Theorem was used to calculate the combined angles (Victor, Harbluk & Engstrom, 2005), i.e. the combined angle is the square root of the sum of squared vertical and squared horizontal angles.

$$\text{Gaze Angle} = \sqrt{(\text{Vertical Angle})^2 + (\text{Horizontal Angle})^2} \qquad (7.10)$$

A decreased SD Gaze suggests a shrunk searching range, less active searching behaviour and potentially an omission of peripheral vision, as gazes are more likely to be located in the central gaze area (i.e. Front).

When performing visual tasks, it is expected that higher SD Gaze would appear, mostly caused by looking at the touch screen, i.e. more gazes looking away from the centre will result in inflation in this measurement. Therefore, the effect of potentially narrowed attention will be dominated by the effect of looking at the touch screen alone. In order to investigate the primary driving related visual dispersion, when performing visual tasks, the SD Gaze for gaze angles looking at Front (naming SD Gaze_Front) was calculated and analysed instead of gaze for the whole area. Similarly, a decreased SD Gaze_Front indicates a shrunk searching range and less active searching behaviour in the Front area.

7.3.3 Blink Rate and Duration

Blink Rate is the number of blinks per minute, and Blink Duration is the duration of each blink (in second). Blink rate is proven to be sensitive to both mental and visual demands (Recarte et al., 2008). Blink rate also was reported to be a predictive and reliable variable for indicating mental workload (Fogarty & Stern, 1989; Recarte et al., 2008; van Orden et al., 2001; Wilson et al., 1994). A higher blink rate is an indication of higher workload, especially for mental workload without visual interference (see Chapter 2). Blink duration was used as another measurement to assess the workload changes, because it is not only related to mental workload, but also to visual workload demands. Neumann (Neumann & Lipp, 2002) suggested it is because decreased blink duration will reduce the interruptions of visual input when the eye lids are closed during a blink. Results from previous research have revealed that blink duration was reduced in actual versus simulated flight (Wilson, Fullenkamp & Davis, 1994).

7.3.4 Saccade Rate, Duration, Velocity, and Amplitude

Saccade Rate is the number of saccades that drivers make per minute, and Saccade Duration is the duration of each saccade in second. Under conditions of increased cognitive load, drivers made fewer saccades (Harbluk, Noy & Eizenman, 2002). Lower saccade rate was consistent with finding of the reduction in glance frequency and less exploration of the driving environment (Recarte & Nunes, 2000). Therefore, in auditory tasking, lower saccade rate is used as an indication of high mental engagement; while in visual tasking, a higher saccade rate suggests a more active searching and an effort applied when perform visual task whilst driving.

Peak of saccadic velocity was suggested to be a good indicator of subjective mental workload, especially for the risk-prone drivers (Di Stasi et al., 2009). A negative correlation of peak of saccadic velocity and subjective mental workload was found for high risk-prone drivers, which suggests in this case, higher saccade velocity waw associated with a lower level of subjective workload. When risk-proneness was low, the opposite result was observed. The saccade velocity was computed by means of a two-point central difference (Larsson, 2002):

$$\partial y(\mathbf{x}) = \frac{y(\mathbf{x}+h) - y(\mathbf{x}-h)}{2h} \qquad (7.11)$$

applied to each gaze component and then weighted together with a square-sum-root,

$$v = \sqrt{\partial y_{yaw}^2 + \partial y_{pitch}^2} \qquad (7.12)$$

Based on the previous research (Di Stasi et al., 2009), it was expected that higher saccadic velocity for the risk-prone drivers (in this study, young and least experienced drivers, see Chapter 6) would be connected with a lower workload; while for less risk-prone drivers, should be related to a higher workload.

Saccade Amplitude describes the dispersion of gaze when a saccade occurs. In a visual searching process, large saccades (saccades with large amplitude) always occur before a long fixation, when information-extraction happens. Again, the combined vertical and horizontal angles were used to describe the gaze dispersion, and the amplitude was the subtraction between the maximum and the minimum values in a saccade. In this study, the higher saccade rate was used as an indication of active searching behaviour, and higher saccade duration and amplitude were indications of expanding searching behaviour, in terms of time and spatial aspect, respectively.

7.3.5 Controlled Visual Behaviour Measurements

In order to eliminate the individual variability, measurements collected in baseline driving were used as a control. The values for each measure were subtracted from the values of the average baseline value for each subject. This method is introduced by Russell (Russell, 1990) cited in (Pastor et al., 2006). These new measurements were later proven generally leading to the same conclusion when the differences between driver groups were also taken into account in the analysis. They were only reported when the effect is different to the original values. These measurements were given notations based on the original ones, with a suffix of "_Ctrl".

7.4 Drivers' Individual Differences in Baseline Driving

7.4.1 Gender Difference

In baseline driving, compared to their female counterparts, male drivers spent more time checking on the control panel and less time on looking Front. In detail, male drivers spent 9.92% of time looking at the in-vehicle panel, and females only spent 1.49% of the time; while for the percentage of time spent on looking at areas in front, male drivers were 76.73% and females were 88.49%. The gender differences are illustrated in Figure 7.3.

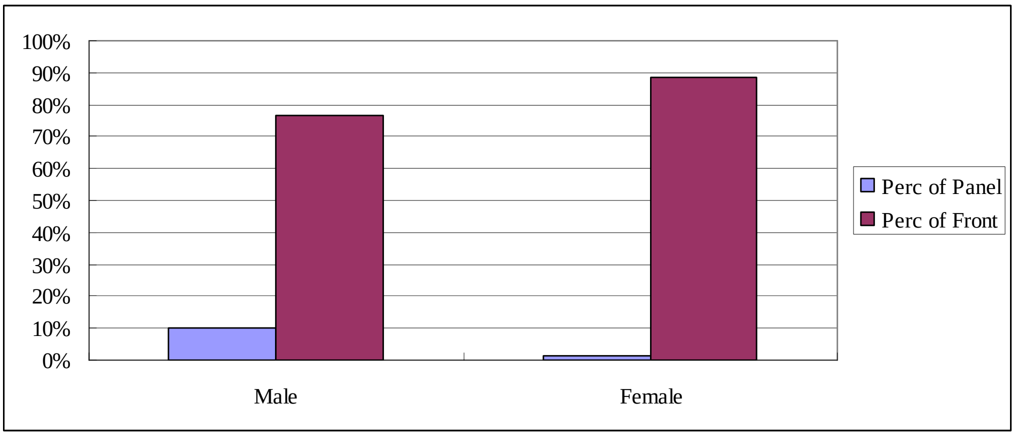


Figure 7.3 Gender difference in visual behaviour during baseline driving

7.4.2 Experience/Age and Annual Mileage

In baseline driving, drivers with more experience (i.e. Experience Group #3) looked less at the road ahead, 71.43% for Group #3 versus 85.76% for Group #1 and 86.23% for Group #2 (Welch Test: $F(2, 67) = 4.104$, $p = 0.021$; BF Test: $F(2, 86) = 6.265$, $p = 0.003$; Tukey HSD, $p < 0.05$). This is consistent with the findings from (Crundall, Underwood & Chapman, 1999a), in which it was explained that experienced drivers are trained to use their peripheral vision for lane keeping and monitoring other on-road information.

The effect of age was similar to that of driving experience, with older drivers (Group #3) showed a higher blink rate (32.84 times per minute versus 27.12 for Group #1 and 24.42 for Group #2, on average, $p < 0.05$), and higher Saccade Rate (24.80 times per minutes versus 11.35 for Group #1 and 11.68 for Group #2). Although the explanation of these age-related spontaneous eye movements is outside the scope of the current study, similar effects were also found in some research on human vision (Sun et al., 1997). The high mileage group also checked more on the panel ($p < 0.01$).

7.5 Impact of Secondary Tasks on Drivers' Visual Behaviour

7.5.1 Drivers' Visual Behaviour in ATCF

7.5.1.1 Effect of Task Levels

SD Gaze and blink rate were significantly impacted by the performance of auditory tasks. The SD Gaze decreased along with the increasing task levels ($F(3, 538) = 2.939$, $p = 0.003$), and the blink rate increased (Brown-Forsythe Test, $F(3, 529) = 4.374$, $p = 0.005$). Post Hoc analysis (SNK and Tukey HSD showed the effects were only significant between the baseline and when performing the auditory tasks). The details are illustrated in Table 7.1. On the whole, performing auditory tasks had no effect on the percentage of time drivers spent on looking at different objects and other eye movement measurements ($p > 0.05$). The reduced SD Gaze suggests when performing auditory tasks, the area of drivers' visual searching shrinks as an effect of high mental workload, known as the "visual tunnel effect". As the effect mainly existed between baseline driving and task performing, but not between the three task difficulty levels, the three levels were combined into one database. From the baseline to tasking the decrease of SD Gaze was from 7.13 to 6.66 degree (decreased by 6.50%), and the increase of blink rate was from 28.52 to 33.82 times per minute (increased by 18.58%).

Table 7.1 Visual behaviour between baseline and auditory tasking in Car-Following

Task Level	0		1		2		3	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
SD Gaze	7.13	1.47	6.73	1.70	6.69	1.68	6.58	1.78
Blink Rate	28.52	14.38	34.41	15.19	32.87	15.78	34.22	16.60

7.5.1.2 Effect of Gender

While performing auditory tasks, the differences between genders from baseline still exist, i.e. male drivers still spent more time checking on the in-vehicle panel, and female drivers spent more time looking ahead ($p < 0.05$). For both male and female, there was an increase in blink rate comparing to baseline driving (Female: ANOVA Test, $F(1, 190) = 6.602$, $p = 0.011$; Male: ANOVA Test, 1,186 = 7.993, $p = 0.005$). Comparison between the two genders when performing auditory tasks, male drivers showed a significantly lower blink rate ($F(1, 396) = 5.199$), $p = 0.023$. The results indicated the workload increased for both genders, and the effect was even more noteworthy for female drivers. These findings are consistent with that found in the subjective rating for workload.

In addition, some more differences between the two genders were observed when performing auditory tasks. Auditory tasks impacted males and females in different ways. Females were found to apply the strategy of checking more often in the mirrors ($F(1, 190) = 4.460$, $p = 0.036$) to deal with the decreased mental resource. For male drivers, on the other hand, the adaptive behaviour of reduced visual searching range was observed (SD Gaze: $p < 0.01$). The reduced gaze variation indicates a less active searching, and may also suggest an omission of the peripheral vision. The details of gender differences when performing auditory tasks comparing with baseline are listed in Table 7.2.

Table 7.2 Gender difference in visual behaviour in ATCF

Gender	Male		Female	
	BL	AT	BL	AT
Panel	9.92%	5.73%	1.49%	2.38%
Front	76.73%	79.52%	88.49%	84.54%
Mirrors	7.34%	6.20%	**5.10%**	**7.59%**
SD Gaze	**7.08**	**6.37**	7.22	7.20
Blink Rate	**27.05**	**32.48**	**31.20**	**36.25**

7.5.1.3 Effect of Experience/Age

Performing the auditory tasks impacted on the three experience groups in different ways. The least experienced group (Driving Experience Group #1) reduced the searching range because of the mental demand caused by auditory tasks, suggested by a significantly lower SD Gaze in the task condition, compared to baseline driving ($F(1, 248) = 9.768$, $p = 0.002$), and an increase of blink rate ($F(1, 248) = 15.700$, $p < 0.001$), suggesting that a higher mental workload has been experienced (Fogarty & Stern, 1989). For Experience Group #2, the only difference found in visual behaviour while performing auditory tasks was the higher blink rate ($F(1, 146) = 5.674$, $p = 0.019$), which also suggested an increasing mental workload. For Experience Group #3, however, no evidence of increased workload was found. Instead, significant increase in the use of mirrors was observed ($F(1, 142) = 4.999$, $p = 0.027$). The details of these visual behaviour changes are presented in Table 7.3.

Table 7.3 Effect of experience on visual behaviour in ATCF

Experience	Group #1		Group #2		Group #3	
	BL	AT	BL	AT	BL	AT
Front (%)	85.76%	87.07%	86.23%	83.07%	71.43%	67.52%
Right Mirrors (%)	1.99%	1.44%	5.06%	4.33%	**4.53%**	**8.47%**
Rear Mirror (%)	5.01%	4.48%	2.32%	2.40%	**2.06%**	**4.78%**
SD Gaze	**7.67**	**7.01**	6.50	6.14	6.88	6.62
Blink Rate	**28.35**	**38.06**	**20.93**	**27.17**	33.73	33.41

From the visual behaviour measurements, it was found that performing the auditory tasks had significantly increased mental workload for the less experienced drivers (Experience Group #1 and #2), reflected by the increased blink rate, but not for the most experienced ones (Experience Group #3). To deal with the extra workload in the dual-tasking condition, the drivers in the least experienced group seem to have adapted their visual behaviour by reducing the searching range, showing a *simplify* strategy, i.e. reducing the intake and processing of information to that which is essential for primary driving. The Experience Group #1 may have reduced the information intake because of the limited central resource (Wickens, 1984b), however, this adaptive behaviour may result in the omission of peripheral vision and other safety-related information. For the experienced driver group, however, another strategy was employed – more mirror checking while performing auditory tasks. The strategy used by Experience Group #3 may be a compensation for obtaining information about on-road traffic more effectively, because more intense information processing is required in the dual-tasking condition, even though this may not be their conscious decision. The higher mirror checking indicated better alertness and higher awareness of traffic situations.

Similar effects were found for the age factor. There was no significant visual behaviour change for Age Group #1. For Age Group #2, a decreased visual searching (lower SD Gaze), increased blink rate and blink duration were observed. For Age Group #3, again, higher mirror checking was found ($p < 0.05$). Again, the factor of experience had a much more pronounced effect in visual searching behaviour than age factor. The details of the age effect when performing auditory tasks are illustrated in Table 7.4.

Table 7.4 Effect of age on visual behaviour in ATCF

Age	Group #1		Group #2		Group #3	
	BL	AT	BL	AT	BL	AT
Front (%)	85.57%	85.56%	87.14%	85.15%	72.13%	71.80%
Rear Mirror (%)	5.46%	5.05%	2.18%	2.29%	**2.50%**	**4.68%**
SD Gaze	7.54	7.37	**6.74**	**6.10**	7.08	6.53
Blink Rate	27.12	32.26	**24.42**	**34.43**	32.84	34.88
Blink Duration	0.16	0.15	**0.16**	**0.17**	0.17	0.17

Similar with that in driving performance, the highest annual mileage group (Mileage Group #3) applied the same copying strategy of checking more into the mirror for a more effective on-road information intake. They increased their time spent on rear mirror from 1.48% in baseline to 4.24% when performing auditory tasks ($p < 0.01$), while this strategy was not applied by Group #1 and #2. Instead, Group #1 significantly reduce their visual searching range (SD Gaze: $F(1, 152) = 6.130$, $p = 0.014$), and Group #2 reduced it marginally ($F(1, 203) = 3.685$, $p = 0.056$), to deal with the high workload, which is also indicated by significant increases in blink rate for these two groups (both $p < 0.05$). Details for the effects of mileage are illustrated in Table 7.5.

Given no significant correlation between the two groups, the similarity of strategies applied by the high driving experienced group and by the more annual mileage group further proved that the coping strategies applied by drivers are trained by practice and are to some extent universal, rather than individual behaviour or only being observed by chance.

Table 7.5 Effect of annual mileage on visual behaviour in ATCF

Mileage	Group # 1		Group # 2		Group # 3	
	BL	AT	BL	AT	BL	AT
SD Gaze	**7.85**	**7.08**	**6.75**	**6.19**	6.88	6.89
Blink Rate	**23.04**	**30.20**	**27.36**	**35.76**	33.78	34.53
Rear Mirror	3.30%	4.47%	5.70%	3.36%	**1.48%**	**4.24%**

7.5.1.4 Effect of Experience of and Attitude towards Using Auditory IVISs

No significant difference in visual behaviour was found between drivers who claimed experience of using mobile phones while driving, and the groups of different attitude towards the in-vehicle auditory systems.

7.5.2 Drivers' Visual Behaviour in VTCF

7.5.2.1 Effect of Task Levels

Performing visual tasks impacted on drivers' visual behaviour dramatically. First of all, significantly more time was spent looking at touch screen (Welch Test: $F(3, 253) = 162.408$, $p < 0.001$; BF Test: $F(3, 420) = 69.729$, $p < 0.001$). Looking at *nothing* was used to indicate *visual transitions* between predefined areas of interest (i.e. front, mirrors, panel and touch screen), which increased significantly when performing visual tasks (Welch Test: $F(3, 291) = 14.032$, $p < 0.001$; BF Test: $F(3, 476) = 8.396$, $p<0.001$). As a consequence of looking on the touch screen and more visual transitions, time spent looking ahead decreased dramatically (Welch Test: $F(3, 303) = 37.416$, $p < 0.001$; BF Test: $F(3, 468) = 56.785$, $p < 0.001$). At the same time, an increase in mirrors checking was found, comparing with baseline driving (Welch Test: $F(3, 301) = 10.354$, $p < 0.001$; BF Test: $F(3, 520) = 8.479$, $p < 0.001$). Consistent with the finding of more visual transitions, when performing visual tasks, the SD Gaze for all gaze angels increased significantly (Welch Test: $F(3, 302) = 32.237$, $p < 0.001$; BF Test: $F(3, 525) = 25.695$, $p < 0.001$). Figure 7.4 gives a visualised illustration of the impact of visual tasks on attention shifted between driving-critical areas.

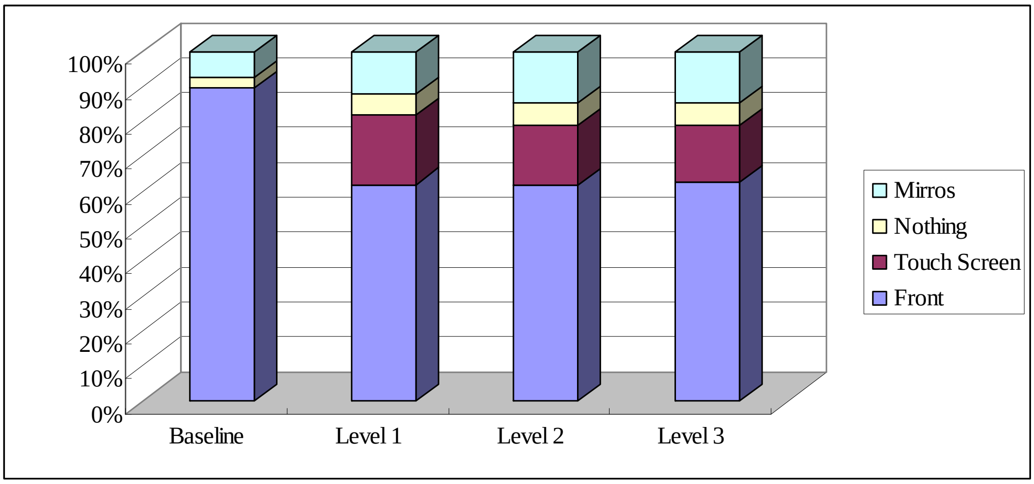


Figure 7.4 The impact of visual tasks on drivers' visual behaviour – by task levels

Apart from the attention shift, more and larger saccade eye movements were also found when performing visual tasks, indicated by higher per minute saccade rate (Welch Test: $F(3, 302) = 18.066$, $p < 0.001$; BF Test: $F(3, 529) = 14.862$, $p < 0.001$), longer saccadic duration (ANOVA Test: $F(3, 488) = 3.094$, $p = 0.027$) and larger saccadic amplitude (ANOVA Test: $F(3, 488) = 8.295$, $p < 0.001$). These changes in drivers' eye movements, again reflected the higher searching demands of the visual task. The impact of visual tasks is illustrated in Table 7.6.

Table 7.6 Visual behaviour between baseline and visual tasking in Car-Following

Task Level	0	1	2	3	VT	Diff (%)
Touch Screen	0%	18.67%	16.31%	15.30%	16.78%	
Front	80.89%	58.26%	57.71%	58.43%	58.14%	-28%
Nothing	2.91%	5.56%	6.03%	5.90%	5.82%	100%
Mirrors	6.54%	11.40%	13.55%	13.67%	12.85%	96%
SD Gaze (°)	7.13	8.62	8.65	8.63	8.63	21%
Saccade Rate	16.58	34.10	34.90	36.17	35.06	111%
Saccade Duration(s)	0.09	0.09	0.10	0.10	0.10	8%
Saccade Amplitude (°)	11.76	15.05	15.17	15.93	15.39	31%

Post Hoc analysis found that all these effects existed only between baseline driving and tasking, apart from the time spent in looking at the touch screen, where significant differences were found between all four levels ($p < 0.05$). The Diff column in Table 7.6 demonstrates the difference between baseline and visual-tasking (average of all three levels) as a percentage, calculated by:

$$\text{Diff} = \frac{\text{Value when performing VT} - \text{Value in BL}}{\text{Value in BL}} \times 100\% \qquad (7.13)$$

7.5.2.2 Effect of Gender

Compared to male drivers, female drivers had a higher blink rate and saccade rate when performing visual tasks ($p < 0.01$), which suggested a higher workload and effort for females. This is consistent with what was found in subjective rating. They paid more attention to mirrors ($p = 0.05$). This might be their strategy to cope with this extra demand, i.e. to obtain traffic information more effectively.

Male drivers, on the other hand, had a much higher percentage of time in visual transitions when dealing with visual tasks (*Nothing*, Welch Test: $F (1, 405) = 37.040$, $p < 0.001$). Higher visual transitions indicated the attention switched between touch screen and road ahead. This result may explain the better secondary task performance of male drivers. As a trade-off, similar with that in auditory tasking, male drivers had a lower active searching on the road ahead (SD Gaze_Front, $F (1, 407) = 4.794$, $p = 0.029$), suggesting that they applied a simplified strategy to reduce the on-road information intake, in order to achieve better task performance. The gender differences in visual behaviour during performing visual tasks are demonstrated in Table 7.7.

Table 7.7 Gender difference in visual behaviour in VTCF

Gender	Male	Female	df	F	P-value
Nothing	6.87%	3.62%	1, 405	37.040	0.000
Mirrors	11.79%	15.09%	1, 187	3.637	0.050
SD Gaze_Front	3.13	3.44	1, 407	4.794	0.029
Blink Rate	29.13	33.69	1, 407	8.556	0.004
Saccade Rate	32.04	41.39	1, 407	8.925	0.003

7.5.2.3 Effect of Experience/Age and Mileage

When performing visual tasks, the visual behaviour of the three experience groups were distinctively different, revealing the effort given to the secondary tasks, and also how they managed their resource and attention. Experience Group #1 made the greatest effort of switching between the primary and secondary tasks, which was suggested by the largest percentage of time spent on visual transitions, high SD gaze and saccade rate. This effort also caused a higher workload for them, indicated by the highest blink rate ($p < 0.01$).

Similar to the findings in the driving performance, Experience Group #3 was the least engaged in the visual tasks, suggested by the lowest percentage of time spent on the touch screen ($F (2, 406) = 19.329$, $p < 0.001$). This argument was also supported by the controlled measure of "looking ahead". Even though performing visual tasks caused decrease in looking ahead for all three groups, Experience Group #3 was the least affected (Percentage of Time Spent on

Looking Front_Ctrl, Welch Test, F (2, 229) = 5.200, p = 0.006). This group of drivers made more compensations during the dual-tasking situation, suggested by more checking on mirrors (F (2, 406) = 9.245, p = 0.001). As discussed before, mirror checking in a dual-tasking condition suggests drivers' intention of extracting on-road information more effectively.

For the details of the impacts of visual tasks on drivers' visual behaviour by drivers' experience group, see Table 7.8.

Table 7.8 Effect of experience on visual behaviour in VTCF

Experience	TS	Nothing	Mirrors	SD Gaze	Blink Rate	Saccade Rate
Group #1	18.10%	**7.01%**	10.07%	**9.02**	**35.21**	**42.65**
Group #2	20.14%	5.01%	13.03%	7.51	22.75	25.41
Group #3	**11.04%**	4.70%	**15.33%**	7.01	31.11	32.44

Similar effects have been found for the age factor. When performing visual tasks, Age Group #3 spent least time on checking the touch screen. A compensation for the high workload was also observed, which was suggested by more checking on mirrors ($p < 0.01$). This strategy was also found for this driver group when performing auditory tasks. Age Group #1, again, was found to pay most attention to the visual tasks, suggested by the highest percentage of time spent on looking at touch screen, and highest SD gaze angle, which in turn caused high saccade rate and amplitude ($p < 0.01$).

Similar but less significant effects were found for mileage factor, where the least annual mileage group made the most visual transition, higher SD Gaze, more and large saccade, which all suggest more effort of visual searching, but these effects were not significant ($p > 0.05$). Details of mileage effects are listed in Table 7.9.

Table 7.9 Impacts of visual tasks on drivers' visual behaviour - by mileage groups

Mileage	Nothing	SD Gaze	Sac Rate	Sac Duration	Sac Amp
Group #1	6.05%	9.26	42.78	0.10	15.24
Group #2	5.52%	7.78	24.17	0.10	15.85
Group #3	5.94%	8.96	39.42	0.09	15.05

7.5.2.4 Effect of Experience of, and Attitude towards, Using Manual IVISs

Drivers who had more previous experience in using manual systems (Expe_Visu Group #1) showed a lower mental workload, indicated by a lower blink rate. This again showed that the workload can be reduced by practice, and in consequence familiarisation and automation of secondary tasks. Even though drivers with more IVISs experience actually showed more active searching on touch screen, with Group #1 showing higher percentage of time looking at touch screen, and higher SD Gaze_Front.

Drivers who had less experience in using these systems (Group #2) showed a higher percentage of time looking at mirrors, to compensate for the extra demand for information processing. This may explain why even with less experience in using IVISs, they still managed to keep their driving performance at the same level as Group #1 (See Chapter 6). Compared to Group #1, they made more and smaller saccades when performing the visual tasks, indicated by a higher saccade rate, shorter saccade duration, and smaller amplitudes. This finding suggested that people who are not familiar with the manual manipulation the in-vehicle systems tend to "chunk" the check on the touch screen into smaller pieces. Details of effects of experience are illustrated in Table 7.10.

Table 7.10 Effects of experience of using manual systems on visual behaviour in VTCF

Expe_Visu2	1	2	df	F	p-value
Mirrors	10.47%	**15.08%**	1, 407	10.866	0.001
TS	**19.86%**	13.89%	1, 407	26.172	0.000
SDGaze_Front	**3.44**	3.03	1, 407	9.666	0.002
Blink Rate	**28.15**	32.91	1, 407	10.72	0.001
Sac Rate	33.47	**36.54**	1, 407	10.219	0.002
Sac Duration	0.10	**0.09**	1, 407	3.302	0.070
Sac Amp	16.70	**14.17**	1, 407	10.72	0.001

For drivers' attitude towards using manual IVISs, those who thought that using manual systems is unsafe (Group #3) paid the least time checking the touch screen, suggesting they are more aware of the risk, and therefore behave more conservatively. They also compensated the extra workload by checking more on the mirrors. Those who thought it is safe to use manual systems whilst driving, on the other hand, showed some less cautious visual behaviour. For example, they spent more time looking away (20.35% on the touch screen, versus 9.38% for Group #3), and they made less saccades with long duration and large amplitude, suggesting them looking away and dwelling on the touch screen for a longer duration each time.

7.6 Summary and Discussion

The main findings from this Chapter are:

- When performing the auditory tasks, there was a significant shrink in drivers' visual searching range, as well as a higher blink rate, suggesting a higher workload. As no substantial deterioration in driving performance was found (see Chapter 6), this shrink in searching range showed the added value of conducting eye movement research, which revealed drivers' change in visual behaviour, lack of awareness to their surroundings, and therefore the safety implications of being mental distracted.

- The visual tasks caused even more pronounced visual behaviour changes in drivers. Apart from extra time spent on looking at the touch screen, and less on the road ahead, a higher visual transition between the defined objects and more mirror checking were also observed. This revealed how drivers compensated for the lost of front vision in real driving conditions, and revealed the *extra* demand caused by manual operation of a visual system.
- Drivers' visual behaviour also revealed the different coping strategies adopted by males and females. When performing the auditory tasks, male drivers simplified the dual tasks by reducing their visual searching range, while female drivers compensated by looking more at the mirrors to collect additional traffic-related information, and their eye movements also showed they experienced a higher workload, which is consistent with their subjective rating. Similar strategies were also adopted for the visual tasks, where male drivers reduced their visual searching to the Front, while females spent a higher percentage on mirror-checking.
- Drivers' visual behaviour also showed that the effect of dual-tasking had a higher impact on less experienced drivers. These drivers experienced a higher workload, as suggested by higher blink rates, which is opposite to their subjective rating, and again confirmed that they had a lower awareness of this potential risk. While performing the auditory tasks, these drivers also reduced their active searching to surrounding traffic conditions (with less mirror-checking, and more to the Front); while the most experienced drivers did the opposite, i.e. spent both more time, and selective, mirror checking. With the visual tasks, the least experienced drivers engaged *more* in the secondary tasks, and took significantly longer and more frequent glances to complete them, suggesting they may find in-vehicle visual tasks more demanding than experienced drivers,. In contrast, the most experienced drivers spent significantly lower amounts of time looking at the touch screen, yet still achieved good performance both in driving and the secondary tasks (see Chapters 5 and 6). Their visual behaviour also revealed that they were more aware of the potential risks involved and took action to protect their overall driving.
- Drivers with more experience of in-vehicle systems showed a lower mental workload in dual-tasking, even though they spent more time looking at the touch screen and made more active visual searching, suggesting that practice can improve the performance in visual tasks. Drivers with less manual IVISs experience made more and smaller saccade while performing the visual tasks, suggesting they tried to simplify the tasks by "chunking" the information from the touch screen into smaller segments.
- Drivers' who thought using manual in-vehicle systems was unsafe spent the least time checking the touch screen and more time on mirror-checking, which revealed they were more aware of the risk, and therefore behaved more conservatively. On the other hand, the

visual behaviour of those who thought manual IVISs as safe were less cautious, for example, they spent more time looking away, with longer and larger saccades, suggesting they checked the touch screen (away from the road ahead) for longer durations each time.

Details of these findings are discussed and explained below.

7.6.1 The Effects of Secondary Tasks on Visual Behaviour

In general, all drivers showed a higher blink rate (of 19%) while performing the *auditory* tasks in Car-Following, which suggests that they experienced a higher workload, which is consistent with their subject rating (see Chapter 5), and a similar increase in blink rate was observed in other research due to a higher mental workload (Fogarty & Stern, 1989; Recarte et al., 2008). Although this extra workload did not result in any material deterioration in driving performance (see Chapter 6), the eye movement data revealed that there was a *significant* reduction in drivers' gaze angles by 6.5%, which suggests a *shrink* in their searching range, which is comparable to findings from previous field test research (Engström, Johansson & Östlund, 2005; Victor, Harbluk & Engstrom, 2005).

Drivers' visual behaviour changed more dramatically while performing the *visual* tasks. Their time spent looking at the Front decreased from a baseline of 81% to 58%, the difference largely replaced with checking the touch screen (an average of 17%). These eye movement measurements are consistent with previous research conducted in car-following (Angell et al., 2006), where the percentage of time spent looking at the front was 81% in baseline, but this was reduced to as low as 53% when the task became very demanding. Although other research (Victor, Harbluk & Engstrom, 2005) suggests that the time spent looking at the front could be lower in baseline driving (at c.70%), this was not conducted under car-following conditions, which require drivers to focus more on the vehicle ahead, as reported by Underwood (Crundall, Shenton & Underwood, 2004). The higher time spent looking at the front was also reported in earlier research (Hughes & Cole, 1986), where 85% of fixations were within the region of 6° of the centre of drivers' focus. However, this was conducted by watching a movie of driving scenes on quiet road sections, where the desire to check other surrounding information might be lower.

In this situation, the reduction in time spent looking at the Front is therefore likely to be a direct consequence of the need to perform visual tasks. In turn, the frequency, duration and amplitude of *saccade* for drivers also increased, due to a need for more visual searching between the touch screen and the viewing required for primary driving. Drivers also increased their mirror-checking (from a 7% baseline to 13%), with more *visual transitions* between the different areas of interest (from 3% to 6%), this being the time not looking at any specific object or the cross over time between objects. This increase in mirror-checking was also found in other higher-demand driving situations (de Waard, 1996a), and suggests drivers were aware of the increased workload caused by the visual tasks, and therefore adopted a strategy to compensate for the lost of Front view and less attention on the road by attempting to extract more visual information from the mirrors and to do this more efficiently.

7.6.2 Effects of Driver Characteristics on Visual Behaviour

7.6.2.1 Gender Effects

From the eye movement measurements, female drivers were found to show a higher blink rate than males while performing the *auditory* tasks (36 versus 32 times per minute). This suggests they experienced a higher workload, which is reflected in the difference in their primary driving behaviour (see Chapter 6) and secondary task performance and Subject workload Rating (see Chapter 5). The two genders coped with the extra mental demand from auditory tasks using different strategies, with females choosing to focus on primary driving and therefore compensate by increasing steering wheel adjustments and longer headways, and paying more attention to mirrors in order to extract more visual information. Male drivers did not adapt their driving behaviour, perhaps because they did not experience as high an *increase* in workload *relatively* due to the auditory tasks, as the blink rate suggests, and this accounts for the lack of deterioration in their driving performance. *However*, their eye movement showed their gaze angle deviation reduced by 10%, whereas this stayed constant for females. This significant shrinking in their searching range, known as a *visual-tunnel* effect (Victor, Harbluk & Engstrom, 2005), suggests males attempted to lower their overall workload by *reducing* the information intake, which freed some resources for male drivers to perform the auditory tasks, and this may explain their better performance in these tasks compared to females. While this *simplification* strategy *appeared* effective in the auditory tasks, as there was no deterioration in driving performance, the reduction in visual searching suggests that male drivers may be less aware of the effects of these mental distractions on their visual and driving behaviour (see Chapter 8). Earlier research (Hughes & Cole, 1986) suggest, when drivers' fixations are concentrated on the road centre area, their information intake from the surrounding region could be significantly reduced. This reduction in the critical information needed for driving, and the reduced attention to surroundings, could potentially increase the risk of accidents and lead to safety concerns; and while adopting this strategy may appear to work for low resource-conflict or auditory tasks for male drivers, it certainly did not work for high conflict or visual tasks, where there was significant driving performance deterioration.

For the *visual* tasks, the Front gaze deviation of male drivers is 10% lower than females, although their visual transitions between different areas of interest is higher (by 3%), which indicates they again focussed on a narrower viewing area at the Front, i.e. a smaller on-road visual searching range. While this simplification strategy can reduce information processing as indicated previously, and benefit them by releasing the central resource for dual-tasking (Wickens, 1984b), the strategy still resulted in deteriorations in driving performance, and they needed to make more effort and to compensate for their driving behaviour, which was not as stable as female drivers in both Scenarios (see Chapter 6). Again, through less active searching, the accident risk for male drivers is potentially increased due to a reduced peripheral vision and therefore potential to miss other important or safety-critical on-road information.

A higher blink rate for female drivers in visual tasks (34 times per minute) again suggests they experienced a higher workload than males (29 per minute), and females adopted a similar compensation strategy to performing auditory tasks, by paying more attention to mirrors to extract more information (a 10% increase over baseline) and do this more effectively (their saccade increased by 29 times per minute over baseline). This compares to only a minor increase in mirrors (4%) for male drivers and a lower increase in saccade (13 times per minute), which again suggests females drive more attentively and devote more effort than males (Dorn et al., 1991), which result in more stable vehicle control generally.

7.6.2.2 Effect of Driving Experience/Age/Mileage

In terms of visual behaviour, dual-tasking had lower impacts on the more experienced drivers' Groups (#2 and 3). The Group with the least experience (#1) showed a higher blink rate, which suggests a higher workload on these drivers than on the more experienced Groups. When performing the auditory tasks, Group #1 increased their percentage viewing to the Front, and reduced the look to mirrors, which suggests an effect of visual-tunnelling (as described above) and less active searching of surrounding traffic conditions; while the most experienced drivers (Group #3) did the opposite, and spent both more time on mirrors and more selective mirror checking (e.g. using the right-mirror in addition to rear mirror), which suggests increased awareness of the additional task demand and alertness to the need for extra vigilance in Car-Following, and a similar difference in mirror checking between different experience groups was reported by Underwood et al. (Underwood, Crundall & Chapman, 2002).

The least experienced drivers also made more saccades (43 times a minute versus 32 for Group #3), which suggests they paid more attention on searching forwards and backwards between the touch screen and the road-ahead, despite experiencing a higher workload compared to the other groups, as indicated by a higher blink rate. This shows an intention to perform well in the visual tasks, and therefore their lower secondary and primary task performance compared to other Groups (see Chapters 5 and 6 respectively) is purely due to a lower awareness of the extra workload involved in dual-tasking and a lower capacity to process the information needed, which is consistent with previous studies (Lansdown, 2002; Wikman et al., 1998), which suggest less experienced drivers find in-vehicle visual tasks more demanding than experienced ones, and they took significantly longer time and more frequent glances to complete tasks.

It is also clear from the eye movement data that the most experienced drivers *intentionally* engaged less in performing the visual tasks - only 11% of their time was spent looking at the touch screen, as opposed to 15% on mirrors, and 5% on visual transitions (or nothing). This compares to 18% for the touch screen for Group #1, and 10% on mirrors, and 7% on visual transitions. The less time spent viewing the touch screen suggests the most experienced drivers adopted a different coping strategy, which de-prioritises visual tasks over primary driving, so they are not easily distracted, and this resulted in a similar secondary task and driving performance (see Chapters 5 and 6) as the next experienced Group (#2), who spent more time on the touch screen (20%) and less on mirrors (13%). By paying more attention to mirrors, these drivers achieved considerably higher awareness and alertness of their surroundings, and therefore the best possible balance between safety and task performance whilst driving, although this did not *necessarily* result in the best primary or secondary task performance metric in every case. This different coping strategy allowed them to allocate their resource and attention not only to the primary and secondary tasks, but also to process more information which is essential to safe driving when the potential risks are increased through dual-tasking. Similar effects were observed for the different age and annual mileage groups, as referred to in the secondary and primary task performance discussions in Chapters 5 and 6.

These overall performance differences between the three experience/age/mileage Groups is consistent with the *skill-rule-knowledge* theory (Rasmussen, 1987) or *control-manoeuvring-strategy* in driving (Michon, 1985) - see Chapter 2, which states that skill or *control* is the first and easiest stage to be obtained. In this case, control represents both driving and performing secondary tasks, and drivers can be trained to be skilled at this level through many years of experience, as was the case with Group #2, which is reflected in their higher performance in dual-tasking compared to less experienced drivers in Group #1. However, the *higher* behaviour levels of rule or *manoeuvring* and knowledge or *strategy* can only be gained through *even more* experience, and the better drivers are at the control level, the easier it is for them to develop and manage manoeuvring- and strategy-based behaviour, which accounts for the difference between the most experienced drivers in Group #3 and those in Group #2, i.e. the most experienced drivers take an overall strategic view of *all* the risks and requirements involved, and *optimises* the balance of outcomes between task performance and safety. This supports Michon's model (Michon, 1985) to suggest that, with significantly increased experience, behaviour at the manoeuvring level extends beyond increasing the capacity for performing tasks in complicated situations, to adopting a more effective strategy in driving. However, this increased awareness of surroundings as a higher-level behaviour has only been observed in this dual-tasking situation, through the use of eye movement measurements.

In conclusion, while an analysis of driving and secondary task performance measurements (from Chapters 5 and 6) reveal that *more* experienced drivers do perform better at the control level, they failed to show any differentiation with *further* experience, as the performance of drivers in Groups #2 and 3 were similar, and did not suggest that the most experienced drivers achieved a *higher* manoeuvring and strategy level of behaviour. However, investigations of drivers' visual behaviour *do* suggest they achieved a higher state of performance, by being more aware and alert to their surroundings, which only becomes apparent through their eye movement.

7.6.2.3 Effect of Experience/Attitude towards Using IVISs

Experience in using auditory IVISs whilst driving did not impact drivers' visual behaviour. However, drivers who had more experience in manual IVISs, or Expe_Visu Group #1, showed a lower mental workload, as indicated by a lower brink rate (28 times per minutes versus 33 for Group #2), even though they spent more time looking at the touch screen (20% versus 14%) and made more active visual searching (a larger searching range for the Front than Group #2). Therefore practice can improve the performance of visual tasks, due to familiarisation and automation in this skill-level, as discussed before. This in turn reduced the overall workload for these drivers when dual-tasking.

Drivers who had less experience of manual IVIS (Expe_Visu Group #2) therefore had relatively higher workload, and they looked more at mirrors (15.08% versus 10.47%) to compensate, which is also reflected in their driving behaviour (See Chapter 6), i.e. they slowed down or increased headway to maintain the same lane control, and this explains their driving performance comparable to the other Group (#1). However, this was at the expense of accuracy in visual task performance (see Chapter 5), i.e. they were either less trained with performing the visual tasks or they prioritised their resources to driving. Drivers with less manual IVISs experience also made more and smaller saccade while performing the visual tasks. They showed a higher saccade rate compared to Group #1 (36.54 times per minutes, versus 33.47 on average), a shorter saccade duration (0.09s versus 0.1s), and smaller amplitude (14.17°versus 16.7), which suggest they tend to "chunk" the information intake from the touch screen into smaller segments. A study of visual behaviour can therefore also reveal how drivers organise their information.

Drivers' who thought using manual systems is unsafe (Atti_Visu Group #3) spent the least time checking on touch screen (9.38% versus 20.35% for Group #1), suggesting they were more aware of the risk, and therefore behave more conservatively. This is also supported by more mirror checking (22.23% versus 11.77% for the Group #1) and conservative primary driving (see Chapter 6). These less looking away, more checking on the mirrors, compensation in lower speed in turn led to faster reaction time in secondary tasks, and resulted in a lower workload (lower subjective rating, see Chapter 5). Those who thought manual IVISs as safe (Atti_Visu Group #1), on the other hand, showed less cautious visual behaviour. For example, they spent more time looking away, made less saccades (30 times per minute versus 47 for Group #3) but with longer duration (0.1s versus 0.09s for Group #3) and large amplitude (16.54°versus 13.40 for Group #3), suggesting them looking away and checking the touch screen for a longer duration each time. Combining the information of more time spent on checking touch screen, looking away for longer duration, less compensation in primary driving (See Chapter 6), and consequently higher workload (See Chapter 5), these results suggest that this group of drivers may leave themselves in a more demanding situation, and result in a more risky condition over time, for example, in Free-Driving, a deterioration in lane control was found. In summary, a study in drivers' visual behaviour differences between different experience groups of using IVISs revealed how workload varies and how information was organised due to various experience; while the comparison between attitude towards IVISs revealed more about the cautious level and different coping strategies between different attitude groups.

Chapter 8 Drivers' Self-Perception on Driving Performance and Visual Behaviour

8.1 Introduction

This chapter describes the results of the post-driving interview on drivers' self-perception of their driving behaviour changes when performing secondary tasks. These self-perceptions were compared against their actual behaviour in order to understand the relations and gaps between the two. The effects of drivers' characteristics, including age, gender, driving experience, annual mileage, experience and attitude towards using IVISs, on these relations and gaps were investigated and are presented in this Chapter.

The main objectives for exploring drivers' self-perception are:

- To understand the drivers' self-awareness of their behaviour changes;
- To study the impact of driver characteristics on their self-awareness;
- To identify the relationship and gap between the actual behaviour and self-claimed behaviour.

The interviews were conducted immediately after the test runs (for detail, see Chapter 3). The aim of the interview was to collect the drivers' self-awareness of their behaviour changes, and their opinions about the experience of being mentally and visually distracted whilst driving.

8.2 The Post-Driving Interview

8.2.1 The Structure of the Interview

The interview was composed of three sections (see Appendix C for more details): Sections A, B and C. Section A was to compare the drivers' experienced workload on the whole, when performing two different types of secondary tasks in two scenarios of Car-Following and Free-Driving. In this section, subjects were required to rate the demands for each type of task in each driving scenario by comparing them with "normal driving", on a scale from 1 (representing "no more demanding") to 5 (denoting "most demanding"). Sections B and C are related to the detailed behaviour changes while performing these tasks. Seven hypotheses on behaviour changes were listed, and participants were asked if they found their behaviour changed in a way confirming the hypothesis or not. The hypotheses are, when performing secondary tasks, drivers:

1. reduced speed;

2. increased headway;
3. started to reduce speed or increase headway before the tasks had started;
4. made more steering wheel adjustments;
5. looked less at adjacent traffic;
6. reacted more slowly;
7. made other self-aware behaviour changes.

And the options of answers were "Yes", "No", "Sometimes" and "Against". In the end, the participants were asked if the behaviour changes were conscious decisions or adaptive reactions towards the dual-tasking.

8.2.2 Participants

Same subjects who participated in the test runs were interviewed (as mentioned in Chapter 5), however, 7 out of 41 of the datasets were incomplete, and therefore data collected for these subjects were excluded in this analysis. The following analysis is based on the interview results from the remaining 34 subjects.

8.3 Results of Post-Driving Interview

8.3.1 Section A

According to the Post-Driving interview, there was no significant difference between the workload of the auditory and visual tasks (see Figure 8.1), where AT and VT stand for performing auditory and visual tasks respectively, while CF and FD stand for in car-following and free-driving scenario respectively.. However, there was a difference between the two driving scenarios. The Car-Following was found to be significantly more demanding as a driving scenario than the Free-Driving for both types of secondary tasks (ANOVA Test and Tukey HSD, $p < 0.001$).

While following a vehicle ahead, in addition to dual-tasking, drivers need to process the information from the vehicle straight ahead, to avoid risk, which explains why they have experienced a higher workload in Car-Following than Free-Driving. Previous research (Crundall, Shenton & Underwood, 2004) also showed that in an intentional car following, driving performance deteriorated, drivers made more errors and more accidents in a study conducted by a simulator. Also, the visual searching range was reduced and information processing became more demanding, which was related to a longer time spent on visually processing the vehicle ahead.

The higher workload in Car-Following was however not reflected in the subjective rating of each specific task conducted during the experiment. This may suggest that during a demanding task condition (dual-tasking in this case), drivers could not differentiate the smaller workload difference due to the background driving condition. This difference between driving scenarios was only realised after the test run, and reflected in the post-driving interview.

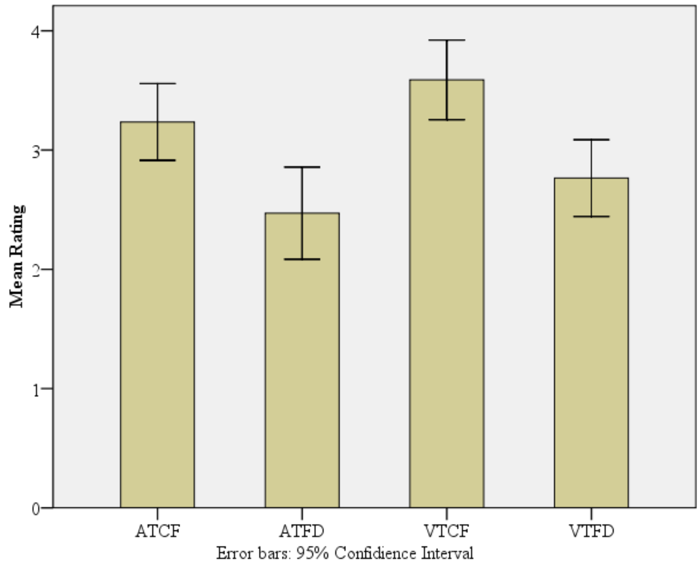


Figure 8.1 Ratings on task demands

8.3.2 Section B (Auditory Task)

Most drivers confirmed the hypothesis of behaviour changes by admitting that they had changed their driving behaviours in some way, when performing the auditory task. Only 3 drivers out of 34 claimed that their behaviour was not affected in any way by performing auditory tasks. All these three drivers were males, and from age groups #2 and #3.

Only 6 participants out of 34 (18%) claimed that they had changed their behaviour before the tasks had started, in order to get themselves ready for the dual-tasking, most of them said they never did so (20, 59%). 17 participants (50%) deemed that they did not change their steering behaviour, and 15 (44%) did not think they changed their visual behaviour. More than 65% (23 participants in total) indicated that they have, or sometimes have, reacted slowly when performing auditory tasks. Drivers' ideas about speed change seemed to be less certain, as the majority answering to "reduced speed" was "sometimes" (70.59%). For the detailed drivers' self-perception of behaviour changes, see Table 8.1.

Table 8.1 Drivers' self-percept of their behaviour changes during auditory tasks

Drivers' Self-Perception of Behaviour Changes								
Auditory Tasking	Yes		Sometimes		No		Against	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Reduce Speed	4	11.76%	24	70.59%	6	17.65%	0	0.00%
Increase Headway	6	17.65%	15	44.12%	12	35.29%	1	2.94%
More Steering	4	11.76%	8	23.53%	17	50.00%	5	14.71%
Look Less	9	26.47%	9	26.47%	15	44.12%	1	2.94%
Change Beforehand	6	17.65%	8	23.53%	20	58.82%	0	0.00%
Slow Reaction	5	14.71%	17	50.00%	12	35.29%	0	0.00%

8.3.3 Section C (Visual Task)

Drivers stated more awareness of their behaviour changes when performing visual tasks than that of auditory tasks. All drivers admitted that their behaviour changed to some extent while performing the visual tasks, as illustrated in Table 8.2. For the anticipation behaviour in dual-tasking, 35.29% drivers (12 participants) agreed that they had changed their driving behaviours before the tasks started in order to get themselves ready for the increasing demand and 50% of drivers (17) said they were aware of their slow reaction. These are much higher than that for auditory tasking.

The most self-aware behaviour change caused by performing the visual task was “look less at the adjacent traffic”. Out of the 34 subjects, 18 (52.94%) answered “yes” for “looking less at the traffic” while performing visual tasks. Once the answers of “sometimes” were included, the percentages increase to 79.4% (27 participants). Again, drivers' ideas about speed seemed to be less certain, as the majority answering to “reduced speed” is “sometimes” (55.88%). In total, 97.1% of the participants admitted that they did, or sometimes did, reduce their speed, 91.2% drivers said they may have shorter fixation, and 88.2% increased headway. The assumptions of behaviour changes that drivers denied mostly were: “more steering movement” (32.35%), followed by “slow reaction” (20.59%).

Table 8.2 Drivers' self-percept of their behaviour changes during visual tasks

Visual Tasking	Yes		Sometimes		No		Against	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Reduce Speed	14	41.18%	19	55.88%	1	2.94%	0	0.00%
Increase Headway	15	44.12%	15	44.12%	3	8.82%	1	2.94%
More Steering	14	41.18%	7	20.59%	11	32.35%	2	5.88%
Look Less	18	52.94%	9	26.47%	5	14.71%	2	5.88%
Change Beforehand	12	35.29%	9	26.47%	13	38.24%	0	0.00%
Slow Reaction	17	50.00%	10	29.41%	7	20.59%	0	0.00%

8.4 Relationships between Actual Behaviour and Self-Perception

In order to identify the relationships and gaps between the actual behaviour and driver self-perception, some behaviour changes confirmed from the interview were compared against the measurements collected from actual driving performance. These measurements are:

1. MNSP (Mean Speed): compare with the answer of the question "did you reduce speed when performing auditory/visual tasks?";
2. MNHW (Mean Headway): compare with the answer of the question "did you reduce speed when performing auditory/visual tasks?";
3. RR1.5 and RR3 (time of Steering Wheel Turns larger than 1.5 /3 degrees per minute), compare with the answer of the question "did you increase steering wheel turns when performing auditory/visual tasks?"
4. SD Gaze (Standard Deviation of Gaze Angle): compare with the answer of the question "did you look less on the peripheral area when performing auditory tasks?", and SD Gaze_Front (Standard Deviation of Gaze Angle_Front only) was used when investigating the impact of visual tasks.

Drivers' actual behaviour changes were investigated for each driver individually, by comparing the driving and visual behaviour during performing secondary tasks with that in baseline driving. This actual behaviour change was then compared with drivers' self-perception, in order to find out the relations and gaps between the two. This Section presents the results and analysis of the effect of drivers' characteristics on this relations and gaps. The gap between behaviour and self-perception includes two types of *false perception*:

1. driving behaviour significantly changed, but drivers did not realise it, recorded as *under-report* false perception;
2. a behaviour change was reported, when where was no evidence for it, i.e. over estimate the effect of secondary tasks, recorded as *over-report* false perception.

8.4.1 Auditory Task

For the effects of auditory tasks, drivers were generally aware of their driving performance changes, and most of time, over-estimated the effects. For example only 14.71% drivers had reduced speed when performing auditory tasks, but more than 82% drivers believed they have, or sometimes have. This over-report of the secondary tasks' impact was found for all driving performance measurements. However, it is noteworthy that the visual behaviour change, on the opposite, was not fully recognised. Even though only 29.41% drivers kept their searching behaviour to the same extent as that in baseline driving, 44.12% believed that they did not look less during performing auditory task. Details of actual behaviour changes and drivers' self-perception are listed in Table 8.3. In the table, the column "SP" denotes drivers' Self-Perception, i.e. the percentage of drivers who answered "Yes", "Sometimes", "No" and "Against" the hypothesis, while the column "DP" indicates the percentage of their actual driving performance change.

Table 8.3 Actual driving behaviour vs. self-perception during auditory tasks

	Reduce Speed		Increase Headway		More Steering		Look Less	
	SP	DP	SP	DP	SP	DP	SP	DP
Yes	11.76%	14.71%	17.65%	29.41%	11.76%	20.59%	26.47%	61.76%
Sometimes	70.59%	0.00%	44.12%	0.00%	23.53%	0.00%	26.47%	0.00%
No	17.65%	85.29%	35.29%	64.71%	50.00%	76.47%	44.12%	29.41%
Against	0.00%	0.00%	2.94%	5.88%	14.71%	2.94%	2.94%	8.82%

8.4.2 Visual Task

Drivers tended to *over-report* their behaviour changes for the effects of the visual tasks. All drivers reported that their driving behaviour changed or sometime changed when performing visual tasks. This means all of the subjects were aware of the impact of visual tasks on their driving performance, but again, the effect was over-estimated. This is particularly noteworthy for the speed control, where most of them did not slow down (88.24%), but only one driver (2.94%) deemed that he kept the same speed with baseline driving. Drivers were more aware of the visual behaviour change caused by visual demand, with over-reporting the visual behaviour changes. They denied most of the steering behaviour change. The details are illustrated in Table 8.4.

Table 8.4 Actual driving behaviour vs. self-perception during visual tasks

Visual Tasking	Reduce Speed		Increase Headway		More Steering		Look Less	
	SP	DP	SP	DP	SP	DP	SP	DP
Yes	41.18%	11.76%	44.12%	35.29%	41.18%	58.82%	52.94%	50.00%
Sometimes	55.88%	0.00%	44.12%	0.00%	20.59%	0.00%	26.47%	0.00%
No	2.94%	88.24%	8.82%	64.71%	32.35%	35.29%	14.71%	20.59%
Against	0.00%	0.00%	2.94%	0.00%	5.88%	5.88%	5.88%	29.41%

8.4.3 Comparison between Auditory and Visual Task

It can be observed from Table 8.3 and Table 8.4 that the percentage of drivers who were aware of behaviour changes when performing the visual tasks is higher than that when performing the auditory tasks, which is generally consistent with the findings in driving performance (see

Chapter 6). More drivers said that they had anticipated and adopted the strategy of adapting the behaviour changes before the tasks started, when performing the visual tasks.

Because drivers mostly over-estimated the effect caused by secondary tasks, it was interesting to understand the correct rate of self-perception, especially where behaviour changes existed but drivers were not aware of. The correct and false perception from each individual in different conditions was calculated. Comparing drivers' self-perception when performing auditory and visual tasks, it was found that drivers' capability to correctly percept their own behaviour changes was affected by the task type ($Chi^2 = 6.984$, $p = 0.030$). They are more likely to correctly perceive their behaviour changes when performing auditory tasks than visual tasks. In detail, 71.57% drivers correctly reported their behaviour when performing auditory tasks, while only 57.84% drivers did so when performing visual tasks. Also, they were more likely to over-report behaviour changes during visual tasks. The comparison of percentages of drivers' false perception between auditory and visual tasking is illustrated in Table 8.5.

Table 8.5 Drivers self-perception compared between auditory and visual tasking

	Auditory Tasking	Visual Tasking
Correct Awareness	71.57%	57.84%
Under-Report	12.75%	10.78%
Over-Report	15.69%	31.37%

8.4.4 Effects of Driver Characteristics

8.4.4.1 Gender Difference

The results from driving performance suggested that on the whole, male drivers' performance was less affected by the auditory task performing (See Chapter 6). This difference between genders was also reflected in the post-interview. In the interview, a higher percentage of female drivers admitted that they adapted their driving behaviour by reducing speed, increasing headway and making more steering wheel turns, while more male drivers said they were not aware of such behaviour changes.

For all self-perception driving performance measurements, when compared against the actual behaviour, female and male drivers make the correct perception of their driving behaviour changes to the same level (66.30% and 62.50% for male and female respectively). But for the incorrect perception, comparing between the genders, it was found that generally, male drivers tend to be less aware of, or under-report the impact of secondary tasks on their behaviour

(15.80%), while female drivers tended to over-report it (259.50%), $Chi^2 = 6.573$, $p = 0.037$. Details of the false perception comparison between males and females are listed in Table 8.6.

Table 8.6 Drivers' self-perception compared between male and female

	Male	Female
Correct Awareness	66.30%	62.50%
Under-Report	15.80%	8.00%
Over-Report	17.90%	29.50%

For example, this gender difference can be observed from the awareness of visual behaviour when performing auditory tasks (The figures are listed in Table 8.7). 65.22% of male drivers reduced their visual searching, but only 43.39% of them were aware of this change. Female drivers, on the other hand, when only 54.55% looked less around, 72.72% of them said they did, or sometimes did so.

Table 8.7 Actual visual Behaviour vs. self-perception during auditory tasks compared between genders

Look Less	Male		Female	
	DP	SP	DP	SP
Yes	65.22%	17.39%	54.55%	45.45%
Sometimes		26.09%		27.27%
No	30.43%	56.52%	27.27%	18.18%
Against	4.35%	0%	18.18%	9.09%

8.4.4.2 Experience/Age and Mileage

The less experienced driver group (Experience Group #1) tended to be less aware of, or under-reported their behaviour changes than the other two groups. Especially, the impact of performing auditory tasks was not fully recognised by them ($Chi^2 (4) = 9.865$, $p = 0.043$), as shown in Table 8.8. It also may be seen from the table, that experienced drivers had a better awareness of their behaviour changes.

Table 8.8 False self-perception compared between experience groups in auditory tasks

Experience Group	Correct Awareness		Under-Report		Over-Report	
	Number	Percentage	Number	Percentage	Number	Percentage
# 1	41	68.33%	12	**20%**	7	11.67%
# 2	23	57.50%	6	15%	11	27.50%
# 3	30	83.33%	4	11.11%	2	**5.56%**

Similar effects were found for age and mileage groups, but the effects were not statistically significant, where young drivers showed worse self-perception.

8.4.4.3 Experience of and Attitude towards Using IVISs

Experience of using IVISs had no a significant impact on drivers' self-perception. The attitude towards using auditory systems whilst driving had significant effect on drivers' awareness of driving behaviour (see Table 8.9). The group who thought the systems are safe (Group #1) said their behaviour had changed least, while more drivers from the group who were more against the systems (Group #3) reported that performing auditory tasks had impact on their driving behaviour (Chi^2 (6) = 15.647, p = 0.016). Experience and attitude towards using IVISs had no impact on the accuracy of drivers' self-perception.

Table 8.9 Self-perception of behaviour change by groups of Atti_Audi

	Attitude towards Auditory Systems		
	Group #1	Group #2	Group #3
Yes	6.20%	10.40%	28.60%
Sometimes	34.40%	54.20%	33.90%
No	56.20%	29.20%	32.10%
Against	3.10%	6.20%	5.40%

8.5 Summary and Discussion

The main findings from this Chapter are:

- Post-Driving Interviews provide a good method to compare drivers' perceptions of their performance and what they actually do when experiencing an extra workload or distraction.
- Drivers were largely aware of - and tended to over-estimate - the negative impact of both auditory and visual tasks on their driving performance. While 72% of drivers correctly perceived their behaviour changes while performing the auditory tasks, only 58% did so for the visual tasks. In addition, they generally under-estimated their eye movement changes while performing the auditory tasks, which shows a lack of appreciation for the impact of mental workload on their visual behaviour and awareness.
- The percentage of male and female drivers who correctly perceived behaviour changes were similar. However, male drivers tended to under-report their performance changes, while females tended to over-report them, and this effect was especially noticeable in their eye movements when performing the auditory tasks.
- Drivers with less experience were also less aware or under reported their behaviour changes compared to the more experienced groups. The impact of age and annual mileage showed similar trends but were not as significant.

Details of these findings are discussed and explained below.

Previously, methods to evaluate drivers' self-perception of their behaviour were largely subjective, being based either on opinions from the drivers themselves (Engström, Johansson & Östlund, 2005), or from the opinions of others (Dejoy, 1991; West et al., 1993), or through the use of separate reference groups (Dejoy, 1991). In this study however, drivers' self-perceive behaviour was compared against their actual driving behaviour data obtained through the experiment. Therefore, these results are both subjective in nature and objectively assessed.

In their Post-Driving Interviews, Drivers suggested that the *demand* of performing a secondary task in the Car-Following Scenario was generally higher than in Free-Driving, but there was no difference between the auditory tasks and visual ones. However in their Subjective Rating *during* the experiment, drivers rated the task demand of all the task Scenarios and types as being similar, and this could be because the effect of adding an extra secondary task alone is much larger than that of a Scenario change, and therefore the Subjective Rating was dominated by this effect. The higher demand in Car-Following is supported by previous research (Crundall, Shenton & Underwood, 2004), which showed that drivers' performance deteriorated in this Scenario, and they spent more time in demanding visual searching, and to process information on the vehicle ahead. The difference in demand between the two Scenarios may explain the extra effort applied by female drivers in steering wheel adjustments to keep a more stable headway and lane control. This could also explain the difference in eye movement for male drivers between the two Scenarios, although this difference was not explored.

The Interviews suggest drivers were largely aware of the *impact* of secondary tasks on their driving and visual behaviour. When their actual behaviour was compared against their self-perception in the Interviews, it was found that drivers tend to *over*-estimate these effects, i.e. they said their behaviour changed while performing the secondary tasks when sometimes there was no measured evidence of such. This awareness and over-estimation was even more evident for the visual tasks, for example more drivers said they adapted their driving behaviour to prepare for the visual task than auditory ones (62% versus 41%). However, drivers generally *under*-estimated the impact of performing auditory tasks on their *eye movement*, where less active visual searching was observed (see Chapter 7) than they had diagnosed, which suggests they were less aware of the impact of the *mental* distraction of auditory tasks on their visual behaviour. Nevertheless, drivers showed better self-perception of their *general* behaviour in auditory tasks than for visual ones, with 72% of drivers correctly perceiving their behaviour changes, compared to 58% for visual tasks.

Significant differences were found between self-perception and actual driving behaviour across the driver characteristics Groups. Although the percentage of male and female drivers who correctly perceived a behaviour change was similar (at 66 and 63%), the percentage of male drivers who were less aware or under-reported the impact of secondary tasks was relatively higher at 16% versus 8% for females, who tend to over-report the impact, at 30% versus 18%. This gender difference is consistent with previous research (Dejoy, 1991), where male drivers tended to be more optimistic than females when adjudging their driving skills.

From the Interviews, only 17% of male drivers admitted to looking *less* peripherally while performing the auditory tasks (with 57% categorically saying they *did not*), when in fact the eye movement data showed 65% of them *did*. This compares to 45% of female drivers who admitted to looking less (with only 18% categorically saying they did not), and 55% actually did. Male drivers were therefore less aware of the effect of *auditory* tasks on their performance, perhaps because they felt a lower *increased* workload compared to females (see Chapter 5), and therefore did not compensate their driving behaviour as a consequence. However, eye movement suggests a significant shrinking in their search area (see Chapter 7), although the driving performance data alone showed no deterioration in these metrics (see Chapter 6). This highlights the potential increased accident risk for male drivers while being mentally distracted e.g. through the auditory tasks, which they were seemingly less aware of.

Across the different Experience Groups, drivers with less experience were less aware of or under-reported their behaviour changes compared to the other two groups, for example 20% in the auditory tasks versus 11% for the most experienced drivers, who showed the best awareness (83% correctly perceiving a behaviour change, versus 68% for the least experienced drivers). The impact of age and annual mileage showed similar trends but were not as significant. Similarly, the different experience and attitude in using IVISs Groups had no significant impact on the accuracy of drivers' self-perceptions.

Chapter 9 Conclusions

This study shows the effectiveness of employing eye movement in addition to traditional performance measurements in developing an improved understanding of the relationships between workload and primary and secondary task performance when dual-tasking, and drivers' visual behaviour and coping strategies for dealing with the increased workload.

9.1 Major Findings and Recommendations

In terms of meeting the aims and objectives for this research:

- **This study investigated the effects of secondary in-vehicle tasks on the relationship between workload and performance, which enables an improved understanding of this relationship to be established, along with drivers' visual behaviour and different coping strategies.**

An on-road experiment was conducted to assess drivers' workload, behaviour and task performance in two real-road Scenarios, using two types of additional in-vehicle tasks (auditory and visual) with three levels of difficulty, and across different driver characteristics groups. Previous research (de Waard, 1996a; Meister, 1976) had suggested a *simple* relationship exists between workload and driving performance, i.e. when workload increases initially, performance can stay unaffected up a point (see "Region A" in Figure 2.3 and Figure 2.4, Chapter 2). However, when workload increases above a certain threshold, performance steadily decreases ("Region B" in Figure 2.3 and Figure 2.4), as drivers' resources become limited (Kahneman, 1973; Posner, 1980; Wickens, 1984a) and they cannot invest further effort (Vicente, Thornton & Moray, 1987) to compensate for the increased demand; and when workload reaches beyond an upper limit, their performance will then stay at a minimum ("Region C"), without further deterioration (de Waard, 1996a).

At first glance, the findings from this study would appear to provide quantitative support for the general relationships described in the simplistic models. For example, when performing the auditory tasks, drivers' performance was largely unaffected (i.e. Region A in the simplistic models), which suggest the conflict of resources with primary driving is low, in accordance with Multiple Resource Theory (Wickens, 1984b). However, when the workload increase caused more resource conflicts, as in the case of visual tasks, driving performance did deteriorate, but this decrease was *not* linear, and fluctuations were found across different driver characteristics groups, which also reflected different coping strategies. In addition, there was a greater deterioration in the more difficult visual tasks, which suggests a tailing off in performance which is not a stable minimum. The simplistic models therefore do not explain why performance should be unaffected at a certain level in Regions C, or account for the fluctuations in performance deterioration in Region B due to individual driver characteristics (Cnossen, Brookhuis & Meijman, 1997) or the effects of different driver coping strategies being applied. A detailed analysis of drivers' visual behaviour, as well as traditional driving performance measurements, is therefore required to understand the complex relationship between workload and performance, and the impact of drivers' coping strategies (see below).

- **The impact of different types of in-vehicle tasks on drivers' workload and primary and secondary task performance was investigated, and this provides a more detailed understanding of the relationship between various types of workload on *both* driving and secondary task performance.**

The impact of auditory and visual secondary tasks on drivers' workload and performance are distinctively different. According to their own Subjective Rating (see Chapter 5), drivers' *workload* increased as the level of secondary task difficulty increased, while their *secondary task performance* decreased, as reflected in a reduction in the percentage of correct answers for the auditory tasks, and an increase in reaction time for the visual tasks. This reflected not only the extra workload and therefore effort required in performing the more difficult tasks, but also that drivers may be *unwilling* or *unable* to devote further resources to compensate (Wickens, 1984b) for their secondary task performance in real-road driving conditions.

Driving behaviour was affected generally in both Scenarios by performing the secondary tasks (see Chapter 6), but more significantly for the visual ones. While performing the auditory tasks, there was an increase in steering wheel adjustments and less predictable steering behaviour or higher entropy, although this was not significant in every case. More significant driving performance changes were found for the visual tasks, with again an increase in steering adjustments and entropy, and a higher lane deviation in Car-Following or speed deviation in Free-Driving. The increased steering adjustments and entropy for both task types and Scenarios showed the extra effort for drivers to keep lane control under the higher workload caused by dual-tasking, which suggest drivers were trying to adjust to the higher task demand. However, even with the extra effort applied, there was still some deterioration in lane control in Car-Following or increased speed deviation in Free-Driving. In the case of visual tasks, drivers also chose to either increase their headway (in Car-Following) or slowed down where possible (in Free-Driving) to *compensate* for the higher conflicts of visual and manual resource required.

These changes in driving behaviour were principally observed in dual-tasking compared to baseline driving, but the variation *between* task difficulty levels was not statistically significant, suggesting the workload increase between difficulty levels was less important than *adding* a secondary task whilst driving, which could also be explained by drivers adopting *different* coping strategies across the three task levels. For example, at the higher difficulty levels, drivers may prioritise their resources so that the primary task of driving took precedent over all other tasks, and this accounts for the deterioration in *secondary* task performance as the difficult level increased. In addition, both driving and secondary task performance showed significant variations between different driver characteristics groups, particularly with driving experience and between genders (see below). However, based on an analysis of traditional driving performance measurements (or indeed with secondary task performance measures and drivers' subjective ratings), the driver differences, and the various coping strategies adopted by them, do not immediately become apparent - an investigation of their eye movement is also required.

- **This study shows the impact of increased workload on different driver characteristic groups had significantly different effects on their primary and secondary task performance.**

Previous research (Dejoy, 1991; Ferguson et al., 2007; Harre et al., 1996; Lansdown, 2002) suggest there are significant differences between genders in driving, in particular that female drivers tend to drive more attentively (Dorn et al., 1991), while males tend to drive faster and incur more traffic violations (Reasons et al., 1990). The driving performance measurements in this study generally supported these findings, for example that females tend to keep more stable lateral and longitudinal control (see Chapter 6), which is particularly apparent in the low resource-conflict auditory tasks, while male drivers *appear* not to be significantly affected; and although the driving performance of both male and female drivers were similarly affected by the visual tasks, females were again more conservative in keeping a more stable headway and lower lane deviation. The cause of these different gender reactions to the extra task demand is not however explained in the literature, perhaps because they cannot be explained by subjective opinion or traditional driving performance measurements alone - but an analysis of eye movement in this research has provided some possible explanations (see below).

Previous research (Kass et al., 2007) also suggest that, when confronted with mental distractions, novice drivers possessed less situational awareness and committed more lapses in concentration than their more experienced counterparts, which is reflected in this study, where for example (see Chapter 5), the least experience drivers rated their workload as being lower than the more experienced groups, which suggest they were less aware of the extra workload these tasks involved. However, unlike the previous research, this study also found that *both* the driving and secondary task performances of the least experienced drivers were significantly lower than the more experienced ones, with for example a lower percentage of correct answers for the auditory tasks, and longer reaction times in the visual ones. This would support existing driving behaviour models (Michon, 1985; Rasmussen, 1987), which suggest the more experienced drivers have a higher capacity to compensate for the extra workload involved in dual-tasking (see Chapter 6), as they are able to automate the vehicle control (required at the control or skill level) to release extra resources to improve decision making (Wickens, 1987b) and time sharing (Schumacher, 2001) at the manoeuvring or rule level, which leads to improved driving performance.

However, while the most experienced drivers performed best in baseline (i.e. normal) driving, with lower steering wheel adjustments and entropy, and better lane control than the other groups, the performance between the most and next-experienced groups were broadly similar in dual-tasking, and there appeared to be no significant benefits from further experience, based on the driving performance and secondary task measurements alone (See Chapter 6). It is only through eye movement that the differences between the two higher experienced groups become apparent (see below).

- **This study shows that dual-tasking had a significant effect on drivers' eye movement, which varied according to different characteristics groups.**

In terms of eye movement (see Chapter 7), all drivers showed a higher blink rate and a significant reduction in gaze angles while performing the auditory tasks, which is similar to previous research on mental distractions (Engström, Johansson & Östlund, 2005; Fogarty & Stern, 1989; Recarte et al., 2008; Victor, Harbluk & Engstrom, 2005). For the visual tasks, the eye movement changes were even more dramatic, with significantly increased saccade (or rapid eye movements) and increased visual transitions between different objects. As expected, a significant amount of the time that was spent looking at the Front was replaced by the touch screen, and these findings are consistent with previous research on visual distractions (Angell et al., 2006; Burnett & Joyner, 1997; Victor, Harbluk & Engstrom, 2005). However, in this study, drivers were found to increase their mirror-checking also, which suggest they were aware of the increased workload caused by the visual tasks, and therefore adopted a strategy to *compensate* for the lost of Front view or less attention on the road ahead, by extracting more visual information from the mirrors and to do this more efficiently. This assumption is supported by other research (Fairclough et al., 1993), which state that drivers performed more frequent mirror-checking when they drive through complex (i.e. demanding) road-sections, which involve many vehicles merging and diverging through highway entrances and exits.

Previous research (Faulks, Irwin & Chekaluk, 2009) also suggest females appear more distracted when using a mobile phone than males, although the reason was largely unexplained. However, from the eye movement measurements in this study, female drivers were found to adopt a different coping strategy than males for dealing with the auditory tasks (See Chapter 7). In addition to more conservative driving behaviour (as referred to further above), female drivers also paid more attention to mirrors than males, in order to extract more visual on-road information to focus on driving. Male drivers, on the other hand, did not appear to adapt their driving behaviour, and the auditory tasks did not appear to cause any deterioration in their driving performance. *However*, their eye movement showed a significant shrink in their visual searching range compared to females, which suggest they attempted to lower their overall workload by *reducing* their information intake and processing (Wickens, 1984b), which freed some resources for them to perform the auditory tasks, and may explain their slightly better performance in these tasks. While this *simplification* strategy *appeared* to be effective in the auditory tasks, it did not work for the high resource-conflict or visual ones, where there was significant driving performance deterioration for both genders. The reduction in males' visual searching range also suggests they may be less aware of the impact of mental distractions on their driving and visual behaviour, which is reflected in their Post-Driving Interviews (see Chapter 8). This reduced attention to the surroundings, or a reduction in the critical information needed for driving, could increase their accident risk while dual-tasking and therefore raise road safety concerns.

Eye movement also helps to explain the performance differences between experience groups (see Chapter 7), in particular why the most experienced drivers did not appear to perform better than the next-experienced group, based on traditional driving and secondary task performance measures. For example, the most experienced drivers *intentionally* engaged less in performing the visual tasks than others, by spending the least time on viewing the touch screen, and more time on mirrors, which suggest they adopted a different coping strategy which de-prioritises the visual tasks over primary driving, so they are not easily distracted, and this resulted in higher awareness and alertness to their surroundings, and a similar driving and secondary task performance to the next-experienced group. This finding extends Michon's *control-manoeuvring-strategy* model of driving behaviour (Michon, 1985) to suggest that, with significantly increased experience, drivers' behaviour at the manoeuvring level extends beyond increasing the capacity for performing tasks in complicated situations, to adopting a more effective strategy in driving, which optimises the balance between safety and performance. However, this higher-level behaviour in dual-tasking situation only becomes apparent through eye movement measurements.

- **Drivers' self-perception of their behaviour changes and their actual performance when engaged in the secondary tasks showed agreements as well as gaps, which provides further explanation for the behaviour changes and individual differences. In particular, the gaps highlight drivers' lack of awareness and understanding of their visual behaviour; and understanding these gaps in future would be beneficial for driver training.**

  The Post-Driving Interviews provided a method to directly compare drivers' perceptions of their driving performance and their actual behaviour when encountering an extra workload or distraction. Drivers were largely aware of and tended to over-estimate the negative impact from both auditory and visual tasks. They generally under-estimated their visual behaviour changes, which shows a lack of such understanding and awareness, and this demonstrates the added value of observing drivers' eye movements. Although the percentage of male and female drivers who perceived a behaviour change was similar, male drivers tended to under report their performance change, while female drivers tended to over report them. This effect was especially noticeable for the visual behaviour changes while performing the auditory tasks. Drivers with less experience were less aware of or tend to under report their behaviour changes compared to more experienced ones.

A study of drivers' subjective perception can therefore benefit advanced driver training and coaching. However this cannot replace the need to investigate driving performance, because even where drivers were aware of their behaviour changes, only 72% correctly perceived a change while performing the auditory tasks, and only 58% for the visual ones.

- **This study shows eye movement, when supplementing traditional performance measurements, provides an additional tool to help improve the understanding of the workload-performance relationship in driving, as well as the different coping strategies adopted by drivers in response.**

The findings in this study on the effects of auditory and visual in-vehicle tasks on driving are largely consistent with previous research (see further above). However, by relating drivers' performance with their eye movement, the findings from this study provide a more detailed and complete picture of the effect of an increased workload on drivers' visual behaviour, their attention allocation, how they respond to the secondary tasks, and the impact on their driving performance. For example, eye movement explained the difference between male and female drivers while performing the auditory tasks, which was not explained by traditional driving performance measurements. In particular, it showed male drivers reduced their visual searching range as compensation for these tasks, which was not detected by traditional measures.

In addition, while drivers were shown to be sensitive to changes in workload, these behaviour changes were not fully explained by existing models, and eye movement can be used to highlight (and help explain) the differences between for example driver characteristics groups, and provide an improved understanding of the different coping strategies they adopt, and therefore the impact of these strategies on the relationship between workload imposed and driving performance. For example, the coping strategy adopted by the most experienced drivers, who used their resource most effectively, by engaging significantly less in the visual tasks, was only revealed by their eye movement. It is therefore recommended that eye movement measurements are included in future driving performance measurements.

9.2 Implications and Benefits for Different User Groups

This research has many implications, for example:

- For setting road safety policies, the use of IVISs by young and inexperienced drivers should be limited, as they show significant deteriorations in driving performing when distracted by these in-vehicle tasks, which they do not adequately compensate for. They also showed less awareness of the impact and effect of these distractions on their behaviour and driving performance, which suggest a need for increased education, as well as controlled use.
- In addition, this research shows hand-free systems, i.e. mental tasks, can be as distracting and demanding as those that require manual manipulation. Even though the impact of such tasks on driving is milder, they can nevertheless reduce drivers' awareness and alertness to surroundings, as shown by their eye movements, which has strong safety implications. Existing distraction-related regulations are largely focused on the need for "hands-on-wheel", because of the emphasis on driving performance-related criteria. However, one conclusion from this study shows drivers' visual searching range can be reduced by the mental distractions of IVISs, and these additional risks should be equally considered. For regulation and safety evaluation purposes, eye movement or visual behaviour monitoring should be included as a complimentary, non-intrusive test for the use (and design) of IVIS technologies.
- The inclusion of eye movement measuring tools should also benefit car manufacturers, in-vehicle system developers and end users (i.e. drivers), as current safety monitoring systems are largely focussed on driving performance parameters, for example, lane departure warning.

Although these systems are very useful, with some even incorporating a degree of visual behaviour monitoring, the emphasis has been on fatigue and drowsiness detection, and it would be beneficial to include additional eye movement measurements into current performance-based systems to provide, for example, distraction detection and warning. Such a system could differentiate between sources of distractions by detecting where drivers are looking, and therefore distinguish the mental and visual workloads according to their visual behaviour changes; and by combining this information with driving performance measurements, the system can provide either a warning or additional assistive information to drivers when the distractions exceed certain thresholds, which can improve their safety and help the future development and take-up of in-vehicle systems.

- Finally, this research also showed the different behaviour and coping strategies across various user groups (e.g. gender, experience and attitude), which can be applied to the design and customisation of in-vehicle systems. For example, the design of these systems could be adapted to suit the driving styles and coping strategies of different genders, and those with various attitudes to IVISs.

9.3 Future Research

This study helped to develop a deeper understanding of the relationship between workload and driving performance, which took into account the impact of different driver characteristics and coping strategies. Eye movement was shown to be a successful tool for providing rich information on how their visual and driving behaviour changed, and this Section now discusses some recommendations and direction for future research, which are related to points raised during this study.

Drivers' visual behaviour in the Free-Driving Scenario was not analysed, as the results of their primary and secondary task performance in this scenario was very similar to Car-Following. However, differences between these Scenarios are possible, although a preliminary analysis suggested they would be minor, for example the time that male drivers spent looking at the instrument panel while performing the auditory tasks was higher than females, which may also help explain their ability to maintain better speed control in Free-Driving.

The focus of this research is also on drivers' visual behaviour and attention allocation *within* the vehicle, which helps to develop an understanding of how their information is organised and processed while performing the in-vehicle tasks. However, drivers' attention can also be affected by *extraneously* road and traffic information *outside* the vehicle. For example, during the experiment, it was observed that female drivers tend to pay more attention and longer fixations on road objects, such as an approaching vehicle, in Free-Driving than their male counterparts. However this observation was not tested, and an investigation of drivers' attention distribution on the *differences* between various objects outside the vehicle, for example traffic lights and different approaching vehicles, could further highlight the different behaviour patterns between driver characteristics groups, and provide further explanation and understanding of drivers' capacity and treatment of important road safety information, which could help the design of road signs and markings. Similarly, this research found some differences between the effects of leading and tailing vehicle types on driving behaviour in general. For example, in the Car-Following Scenario, following a van caused a significantly longer headway than *other* vehicle types, while in Free-Driving, less steering adjustments was found when there was a heavy goods vehicle tailing the test vehicle, and eye movement could perhaps reveal the differences or reasons for this behaviour.

In this study, the effects of visual and auditory workload were investigated separately. As discussed in Chapter 6, in a practical IVIS design, it is likely there is a combination of the two types of workload. Therefore, it would be useful to investigate the combined effects of these two.

Research into eye movement, driving performance and the workload caused by IVISs may be used to assist the design and help provide guidance and regulations on their use. However, the research challenge lies in the criteria to be used in deciding which aspects should be allowed (or needs to be controlled), including how these are defined, and the potential risk of accidents they may cause. For example, the driving performance related to the legal limit of alcohol consumption or the use of a hand-held mobile phone can be used as a benchmark for what is not acceptable.

This study can also be conducted on a larger sample, and perhaps in a naturalistic driving scenario (i.e. without an experimenter) in the future to draw a stronger conclusion, especially to highlight the subject differences. This research would require similar in-vehicle design, i.e. mental-oriented workload and separately, a visual intensive interface. With the help of eye monitoring equipment, on-board video systems and synchronised time record in the IVISs, it is possible to distinguish the type of workload and therefore their associated effects on drivers. This will also allow investigation into the longer term effect of IVISs on driving behaviour. The data collected through large scale experiment can be used to develop a driving and visual information database, from which modern data mining techniques can be used to develop criteria for defining the thresholds of performance deterioration, that has been a challenge for researchers and IVISs developers (Brookhuis & De Waard, 2003). Such research on a large scale was not conducted via this study, although it would have been beneficial, for example in investigating the effect of reduced speed and increased headways due to use of IVISs on traffic flows. A possible approach to this question is to apply macro-level simulation, which can be used to control other on-road factors.

Finally, in this experiment, the least experienced drivers showed a higher saccade rate and more visual transitions while performing the visual tasks, which suggest they may have difficulty in spotting the target circles required by these tasks, given they adopt an unsophisticated visual searching strategy (Crundall, Underwood & Chapman, 1999a), which meant the effort to look at the touch screen alone or away from the Front (as require for primary driving) is already higher for them compared to more experienced drivers, and therefore their capacity to resolve different objects on the touch screen *may* be limited. However, this hypothesis was not tested, and could be explored in the future by comparing the reaction time and task accuracy while performing the visual tasks between driving and not, to determine the resource *costs* difference between driver experience groups.

References

Aarts, L. & van Schagen, Ingrid (2006a) Driving speed and the risk of road crashes: A review *Accident Analysis & Prevention*, 38 (2), 215-224.

Aarts, L.T. & van Schagen, I. N. L. G. (2006b) Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention*, 38 (2), 215-224.

Aasman, J., Mulder, G. & Mulder, L.J.M. (1987) Operator effort and the meansurement of heart-rate variability. *Human Factors*, 29, 161-170.

Abdi, H. (2007) The Bonferonni and Šidák Corrections for Multiple Comparisons IN: N. Salkind, ed. *Encyclopedia of Measurement and Statistics*, Thousand Oaks (CA): Sage.

Ahlstrom, U. & Friedman-Berg, F. (2005) *Subjective workload ratings and eye movement activity measures*, Federal Aviation Administration William J. Hughes Technical Ceter.

Alm, H. & Nilsson, L. (1994) Changes in driver behaviour as a function of handsfree mobile phones: A simulator study. *Accident analysis and prevention*, 26 (4), 441-451.

Anderson, N.D., Craik, F.I.M. & Naveh-Benjamin, M. (1998) The attentional demands of encoding and retrieval in younger and older adults: Evidence from divided attention costs. *Psychology and Aging*, 13 (3), 405-423.

Angell, L., Auflick, J., Austria, P.A., Tijerina, L., Biever, W., Diptiman, T., Hogsett, J. & Kiger, S. (2006) *Driver workload metrics project - Task 2 Final Report*, NHTSA.

Anstey, K.J., Wood, J., Lord, S. & Walker, J.G. (2005) Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review*, 25 (1), 45-65.

Anttila, V. & Luoma, J. (2005) Surrogate in-vehicle information systems and driver behaviour in an urban environment: A field study on the effects of visual and cognitive load. *Transportation Research Part F-Traffic Psychology and Behaviour*, 8, 121-133.

Baddeley, A.D. (1996) The concept of working memory IN: S.E. Gathercole, ed. *Models of short-term memory*, Hove, Uk: Psychology Press, 1-27.

Barnett, V. & Lewis, T. (1994) *Outliers in statistical data (3rd ed.)*. New York: Wiley.

Baron, A. & Green, P. (2006) *Safety and usability of speech interfaces for in-vehicle tasks while driving: a brief literature review*, Transportation Research Institute (UMTRI).

Bischoff, D. (2007) *Developing guidelines for managing driver workload and distraction associated with telematic devices*, Washinton, Available from: http://citeseerx.ist.psu.edu.

Blanco, M., Biever, W.J., Gallagher, J.P. & Dingus, T.A. (2006) The impact of secondary task cognitive processing demand on driving performance. *Accident analysis and prevention*, 38, 895-906.

Bland, J.M. & Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Thelancet*, February (8), 307-310.

Boer, E.R. (2000) Behavioral entropy as an index of workload. *Human Factors and Ergonomics Society Annual Meeting Proceedings, Proceedings 3 - Individual Differences*, 4, 125-128.

Boersema, T., Zwaga, H.J.G. & Adams, A.S. (1988) Conspicuity in realistic scenes: An eye-movement measure. *Applied Ergonomics*, 20, 267-273.

Brabyn, J.A., Schneck, M.E., Lott, L.A. & Haegerstrom-Portnoy, G. (2005) Night driving self-restriction: Vision function and gender differences. *Optometry and Vision Science*, 82 (8), 755-764.

Brookhuis, K., De Waard, D. & Mulder, L.J.M. (1994) Measureing driving performance by car-following in traffic. *Ergonomics*, 37, 427-434.

Brookhuis, K.A. (1998) How to measure driving ability under the influence of alcohol and drugs, and why. *Human Psychopharmacology*, 13, S64-S69.

Brookhuis, K.A., De Vries, G. & De Waard, D. (1991) The effects of mobile telephoning on driving performance. *Accident analysis and prevention*, 23 (4), 309-316.

Brookhuis, K.A. & De Waard, D. (2003) On the assessment of criteria for driver impairment: in search of the golden yardstick for driving performance. IN: *the 2th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Park City, Utah, US, July 21-24.

Brookhuis, K.A. & deWaard, D. (2001) Assessment of drivers' workload: Performance and subjective and physiological indexes IN: P.A. Hancock & P.A. Desmond, eds. *Stress, workload, and fatigue*, London: Lawrence Erlbaum Associates.

Brookhuis, K.A., vanWinsum, W., Heijer, T. & Duynstee, M.L. (1999) Assessing behavioural effects of in-vehicle information systems. *Transportation Human Factors*, 1, 261-272.

Brown, I. & Poulton, E.C. (1961a) Some alternative methods for predicting performance among professional drivers in training. *Ergonomics*, 1, 31-40.

Brown, I.D. & Poulton, E.C. (1961b) Measuring the spare ‘mental' capacity of cardrivers by a subsidiary task. *Ergonomics*, 4, 35-40.

Brown, I.D., Ticker, A.H. & Simmonds, D.C. (1969) Interference between concurrent tasks of driving and telephoning. *Journal of Applied Psychology*, 53, 419-424.

Brown, M.B. & Forsythe, A.B. (1974) The small sample behaviour of some statistics which test the equality of several means. *Techometrics*, 16, 385-389.

Buning, H. (1997) Robust analysis of variance. *Journal of Applied Statistics*, 24 (3), 319-332.

Burnett, G. & Joyner, S. (1997) An assessment of moving map and symbol-based route guidance systems IN: Y.I. Noy, ed. *Ergonomics and safety of intelligent driver interfaces*, New Jersey, USA: L. Erlbaum Associates Inc., 115-138.

Burns, P.C., Parkes, A., Burton, S., Smith, R.K. & Burch, D. (2002) *How dangerous is driving with a mobile phone? Benchmarking the impairment to alcohol*.

Caird, J.K., Chisholm, S.L. & Lockhart, J. (2008) Do in-vehicle advanced signs enhance older and younger drivers' intersection performance? Driving simulation and eye movement results. *International Journal of Human-Computer Studies*, 66 (3), 132-144.

Carsten, O.M.J. & Brookhuis, K.A. (2005) Issues arising from the HASTE experiments. *Transportation Research Part F*, 8 (2), 191-196.

Carsten, O.M.J. & Nilsson, L. (2001) Safety assessment of driver assistance systems. *European Journal of Transport and Infrastructure Research*, 1 (3), 225-243.

Chapman, P.R., Underwood, G. & Geoffrey, U. (1998) Visual Search of Dynamic Scenes: Event Types and the Role of Experience in Viewing Driving Situations IN. *Eye Guidance in Reading and Scene Perception*, Amsterdam: Elsevier Science Ltd, 369-393.

Cheng, B., Lin, Q., Song, T., Cui, Y., Wang, L. & Kuzumaki, S. (2011) Analysis of driver brake operation in near-crash situation using naturalistic driving data. *International Journal of Automotive Engineering*, 2 (40), 87-94.

Chiang, D.P., Brooks, A.M. & Weir, D.H.D.H. (2004) On the highway measures of driver glance behavior with an example automobile navigation system. *Applied Ergonomics*, 35 (3), 215-223.

Cnossen, F., Brookhuis, K.A. & Meijman, T. (1997) The effects of in-car information systems on mental workload: A driving simulator study IN: K.A. Brookhuis, D. deWaard & C.M. Weikert, eds. *Simulators and traffic psychology*, Groningen: Centre for Environmental and Traffic Psychology, 151-163.

Cody, D. & Gordon, T. (2007) TRB workshop on driver models: A step towards a comprehensive model of driving? IN: P.C. Cacciabue, ed. *Modelling driver behaviour in automotive environments - Critical issues in driver interactions with intelligent transport systems*, London: Springer, 26-42.

Cole, B.L. & Hughes, P.K. (1984) A field trial of attention and search conspicuity. *Human Factors*, 26, 229-313.

Cooper, P.J. & Zheng, Y. (2002) Turning gap acceptance decision-making: the impact of driver distraction. *Journal of safety research*, 33, 321-335.

Crundall, D., Shenton, C. & Underwood, G. (2004) Eye movements during intentional car following. *Perception*, 33, 975-986.

Crundall, D., Underwood, G. & Chapman, P. (1999a) Driving experience and the functional field of view. *Perception*, 28, 1075-1087.

Crundall, D., Underwood, G. & Chapman, P. (1999b) Peripheral detection rates in drivers IN: A. Gale, ed. *Vision in vehicles*, Amsterdam: Elservier.

de Waard, D. (1996a). *The measurement of drivers' mental workload*, PhD thesis, University of Groningen.

de Waard, D. (1996b). *The measurement of drivers' mental workload*, PhD Thesis thesis, University of Groningen.

Dejoy, D.M. (1991) An examination of gender differences in traffic accident risk perception. *Accident Analysis & Prevention*, 24 (3), 237-246.

Department for Transport, Scottish Government & Welsh Assembly Government 2009, *Reported road casualties Great Britain: 2008 Annual Report*, London: The Stationery Office.

Department for Transport, UK. (2011) *Design Manual for Roads and Bridges (DMRB), Standards for Highways*, Available from: <http://www.standardsforhighways.co.uk/dmrb/>.

deWaard, D. & Brookhuis, K.A. (1997) On the measurement of driver mental workload IN: J.A. Rothengatter & E.C. Vaya, eds. *Traffic and transport psychology: Theory and application*, Oxford: Pergamon.

Di Stasi, L. L., Alvarez-Valbuena, V., Canas, J.J., Maldonado, A., Catena, A., Antoli, A. & Candido, A. (2009) Risk behaviour and mental workload: Multimodal assessment techniques applied to motorbike riding simulation. *Transportation Research Part F*, 12, 361-370.

Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z.R., Jermeland, J. & Knipling, R.R. (2006) *The 100-Car Naturalistic Driving Study*, Virginia Tech Transportation Institute, No. DOT HS 810 593.

Dorn, L., Glendon, A.I., Hoyes, T.W., Matthews, G., Davis, D.R. & Taylor, R.G. (1991) Group differences in driving performance. IN: *Behavioural Research on Road Safety II*, Manchester University.

Duda, R.O. & Hart, P.E. (1973) *Pattern classification and scene analysis*. New York: John Wiley & Sons.

Dunnett, C.W. (1955) A multiple comparison procedure for comparing several treatments with a control. *Journal of American Statistical Association*, 50 (272), 1096-1121.

Edquist, J. (2008). *The effects of visual clutter on driving performance*, PhD thesis, Monash University.

Engel, F. (1971) Visual conspicuity, directed attention and retinal locus. *Vision Research*, 11, 563–576.

Engström, J., Johansson, E. & Östlund, J. (2005) Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8 (2), 97-120.

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

Fairclough, S.H., Brookhuis, K.A. & Vallet, M. (1993) Driver state monitoring system DETER (V2009). IN: *the Advanced Transport Telematics*, Brussels. 330-335, 330-335.

Farber, E. (2000) Visual attention design limits for ITS in-vehicle systems: the Society of Automotive Engineers Standard (AES). IN: *TRB*.

Faulks, I.J., Irwin, J.D. & Chekaluk, E. (2009) Difference in response of male and female drivers to everyday distractions. IN: *Transportation Research Board 88th Annual Meeting*.

Ferguson, S., Teoh, E. & McCartt, A. (2007) Progress in teenage crash risk during the last decade. *Journal of Safety Research*, 38, 137-145.

Fogarty, C. & Stern, J.A. (1989) Eye movements and blinks: Their relationship to higher cognitive processes. *International Journal of Psychophysiology*, 8, 35-42.

Fox, G., Bowden, S. & Smith, M. (1998) On-road assessment of driving competence after brain impairment: Review of current practice and recommendations for a standardized examination. *Archives of Physical Medicine and Rehabilitation*, 79, 1288–1296.

Goodman, M.J., Bents, F.D., Tijerina, L., Wierwille, W.W., Lerner, N. & Benel, D. (1997) *An investigation of the safety implications of wireless communication in vehicles*, Department of Transportation.

Green, P. (1999a) The 15-second rule for driver information systems. IN: *ITS America Ninth Annual Meeting*, Washington, D.C. Intelligent Transportation Society of America.

Green, P. (1999b) Estimating compliance with the 15-second rule for driver-interface usability and safety. IN: *Human Factors and Ergonomics Society 43rd Annual Meeting*, Santa Monica, CA. Human Factors and Ergonomics Society.

Green, P. (2008) Driver interface/HMI standards to minimize driver distraction/overload. IN: *Convergence 2008, Society of Automotive Engineers* Detroit, MI, USA. SAE International.

Griffen, J.A., Rapport, L.J., Bryer, R.C., Bieliauskas, L.A. & Burt, C. (2011) Awareness of Deficits and On-Road Driving Performance. *The Clinical Neuropsychologist*, 25 (7), 1158-1178.

Guensler, R.L. & Willems, P.J. (2002) The role of instrumented vehicle data in transportation decision making. IN: *Decision Making in Urban and Civil Engineering (DMinUCE)*, London, UK. 1-6.

Hancock, P.A., Caird, J.K. & White, H.G. (1990) *The use of simulation for the assessment, training, and testing of older drivers (Technical Report, NIA 90-01)*, National Institute on Aging, National Institutes of Health.

Hancock, P.A. & Parasuraman, R. (1992) Human factors and safety in the design of intelligent vehicle-highway systems (IVHS). *Journal of safety research*, 23, 181-198.

Hancock, P.A. & Warm, J. (1989) A dynamic model of stress and sustained attention. *Human Factors*, 31 (5), 519-537.

Hankey, J.M., Dingus, T.A., Andrews, C., Hanowski, R.J. & Wierwille, W.W. (2000) *In-vehicle information systems behavioral model and design support*, Center for Transportation Research, Virginia Tech.

Harbluk, J. & Lalande, S. (2005) Performing E-mail taks while driving:: The impact of speech-based tasks on visual detection. IN: *Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Rockport, Maine, USA, 27-30, June, 2005. 311-317.

Harbluk, J.L., Noy, Y.I. & Eizenman, M. (2002) *The impact of cognitive distraction on driver visual behaviour and vehicle control*, Ministry of Transport, Canada.

Harbluk, J.L., Noy, Y.I., Trbovich, P.L. & Eizenman, M. (2007a) An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accident analysis and prevention*, 39 (2), 372-379.

Harbluk, J.L., Noy, Y.I., Trbovich, P.L. & Eizenman, M. (2007b) An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accident Analysis & Prevention*, 39 (2), 372-379.

Harms, L. (1991) Variation in drivers' cognitive load: Effects of driving through village areas and rural junctions. *Ergonomics*, 34, 151-160.

Harms, L. & Patter, C. (2003) Peripheral detection as a measure of driver distraction- A study of memory-based versus system-based navigation in a built-up area. *Transportation Research Part F*, 6, 23-36.

Harre, N., Field, J. & Kirkwood, B. (1996) Gender differences and arears of common concern in the driving behaviors and attitudes of adolescents. *Journal of Safety Research*, 27 (3), 163-173.

Hart, S.G. & Staveland, L.E. (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research IN: P.A. Hancock & N. Meshkati, eds. *Human Mental Workload*, Amsterdam: North-Holland, 139-183.

Hayhoe, M. (2000) Vision using routines: a functional account of vision. *VISUAL COGNITION*, 7, 43–64.

He, P. & Kowler, E. (1992) The role of saccades in the perception of texture patterns. *Vision Research*, 32, 2151-2163.

Hilton, A. & Armstrong, R. (2006) Post hoc ANOVA tests. *Microbiologist*, September, 34-36.

Hoffman, J.D., Lee, J.D., McGehee, D.V., Macias, M., Gellatly, A.W. & Trb (2005) Visual sampling of in-vehicle text messages - Effects of number of lines, page presentation, and message control. IN: *84th Annual Meeting of the Transportation-Research-Board*, Washington, DC, Jan 09-13. Transportation Research Board Natl Research Council, 22-30.

Hong, I., Iwasaki, M., Furuichi, T. & Kadoma, T. (2006) Eye movement and driving behaviour in curved section passages of an urban motorway. *Proc. IMechE Part D: J. Automobile Engineering*, 220, 13.

Horberry, T., Anderson, J., Regan, M.A., Triggs, T.J. & Brown, J. (2006) Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident analysis and prevention*, 38 (1), 185-191.

Horrey, W.J. & Lesch, M.F. (2009) Driver-initiated distractions: examining strategic adaptation for in-vehicle task initation. *Accident analysis and prevention*, 41, 115-122.

Hoyos, C.G. (1988) Mental load and risk in traffic behaviour. *Ergonomics*, 31 (4).

Huestegge, L., Skottke, E.-M., Anders, S., Müsseler, J. & Debus, G. (2010) The development of hazard perception: Dissociation of visual orientation and hazard processing. *Transportation Research Part F: Traffic Psychology and Behaviour*, In Press, Corrected Proof.

Huey, E.B. (1908) *The psychology and pedagogy of reading*. New York, USA: Macmillan.

Hughes, P.K. & Cole, B.L. (1986) What attracts attention when driving? *Ergonomics*, 29 (3), 377-390.

Inhoff, A.W., Solomon, M.S., Seymour, B.A. & Radach, R. (2008) Eye position changes during reading fixations are spatially selective. *Vision Research*, 48 (8), 1027-1039.

Iqbal, S., T., Adamczyk, P.D., Zheng, X.S. & Bailey, B.P. (2005) Towards an index of opportunity: understanding changes in mental workload during task execution. IN: *the ACM Conference on Human Factors in Computing Systems*. ACM press, New York, 311-320.

Jahn, G., Oehme, A., Krems, J.F. & Gelau, C. (2005) Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F-Traffic Psychology and Behaviour*, 8, 21.

Jamson, A.H. & Merat, N. (2005) Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. *Transportation Research Part F*, 8 (2), 76-96.

Janssen, W.H., Brouwer, R.F.T. & Huang, Y. (2004) *Risk trade-offs between driving behaviour and driver state*, TNO Human Factors.

Jarrell, M.G. (1994) A comparison of two procedures, the Mahalanobis Distance and the Andrews-Pregibon

Statistic, for identifying multivariate outliers. *Research in the schools*, 1, 49-58.

Just, M.A. & Carpenter, P.A. (1980) A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87 (329-354.).

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

Kaptein , N.A., Theeuwes, J. & Van Der Horst, Richard (1996) Driving simulator validity: Some considerations. *Transportation Research Record: Journal of the Transportation Research Board*, 1550, 30-36.

Kass, S.J., Cole, K.S. & Stanny, C.J. (2007) Effects of distraction and experience on situation awareness and simulated driving. *Transportation Research Part F*, 10, 321–329.

Klockers, A.J. & Sax, G. (1986) Post Hoc comparisons pairwise methods IN. *Multiple Comparisons*, Beverly Hills, California, USA: Sage Publication, 57-58.

Kramer (1991) Physiological metrics of mental workload: a review of recent progress IN: D.L. Damos, ed. *Multiple-task performance*, London: Taylor & Francis, 279-328.

Kua, A., Korner-Bitensky, N., Desrosiers, J., Man-Son-Hing, M. & Marshall, S. (2007) Older driver retraining: A systematic review of evidence of effectiveness. *Journal of safety research*, 38 (1), 81-90.

Kutila, M.H., Jokela, M., Makinen, T., Viitanen, J., Markkula, G. & Victor, T.W. (2007) Driver cognitive distraction detection: feature estimation and implementation. *Proceedings of the Institution of Mechanical Engineers. Part D, Journal of automobile engineering*, 221, 14.

Labiale, G. (1996) Complexity of in-car visual messages and drivers' performance IN: A.G. Gale, I.D. Brown, C.M. Haslegrave & S.P. Taaylor, eds. *Vision in Vehicles-V*, Bournemouth, UK: Elsevier Science B. V., 187-194.

Lamble, D., Laakso, M. & Summala, H. (1999) Detection thresholds in car following situations and peripheral vision: implications for positioning of visually demanding in-car displays. *Ergonomics*, 42 (6), 807-815.

Land, M. & Horwood, J. (1995) Which parts of the road guide steering? *Nature*, 377, 2.

Land, M.F. (1992) Predictable Eye Head Coordination During Driving. *Nature*, 359 (6393), 318-320.

Land, M.F. (1999) Motion and vision: why animals move their eyes. *Journal of Comparative Physiology A 185*, 341-352.

Land, M.F. (2006) Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, 25, 29.

Land, M.F. & Lee, D.N. (1994) Where we look when we steer. *Nature*, 369 (30), 3.

Lansdown, T.C. (2002) Individual differences during driver secondary task performance: verbal protocol and visual allocation findings. *Accident Analysis & Prevention*, 34, 655-662.

Larsson, P. (2002). *Automatic visual behavior analysis*, MSc Thesis thesis, Linkoping University.

Lee, J.D., Caven, B., Haake, S. & Brown, T.L. (2001) Speech-based interaction with in-vehicle computers: The effect of speech-based E-mail on drivers' attention to the roadway. *Human Factors*, 43 (4), 10.

Lee, P.J.N. & Triggs, T.J. (1976) The effect of driving demand and roadway environment on peripheral visual detections. IN: *ARRB*, Australia. 7-12.

Lengenfelder, J., Schultheis, M.T., Al-Shihabi, T., Mourant, R. & DeLuca, J. (2002) Divided Attention and Driving: A Pilot Study Using Virtual Reality Technology. *Journal of Head Trauma Rehabilitation*, 17 (1), 26-37.

Liu, Y.-c. (2001) Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics*, 44 (4), 425-442.

Liu, Y.-C. & Wen, M.-H. (2004) Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. *International Journal of Human-Computer Studies*, 61 (5), 679-697.

Lunenfeld, H. & Alexander, G.J. (1990) *A user's guide to positive guidance (3th Edition)*. Washington, D. C.: Federal Highway Administration.

Ma, R. & Kaber, D.B. (2005) Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, 35 (10), 939-953.

Ma, R. & Kaber, D.B. (2007) Effects of in-vehicle navigation assistance and performance on driver trust and vehicle control. *International Journal of Industrial Ergonomics*, 37, 665-673.

Marple-Horvat, D.E., Chattington, M., Anglesea, M., Ashford, D.G., Wilson, M. & Keil, D. (2005) Prevention of coordinated eye movements and steering impairs performance. *Exp Brian Res*, 163, 411-420.

Martens, M.H. & Van Winsum, W. 2000, *Measuring Distraction: the Peripheral Distraction Task*, *NHTSA Internet Forum on the Safety Impact of Driver Distraction when Using*

In-Vehicle Technologies, <www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/PDF/34.PDF>.

Mattes, S. (2003) The lane-change-task as a tool for driver distraction evaluation IN: H. Strasser, K. Kluth, H. Rausch & H. Bubb, eds. *Quality of work and products in enterprises of the future*, Stuttgart: Ergonomia, 57.

Matthews, R., Legg, S. & Charlton, S. (2003) The effect of cell phone type on drivers subjective workload during concurrent driving and conversing. *Accident analysis and prevention*, 35 (4), 451-457.

Mayhew, D.R., Simpson, H.M., Wood, K.M., Lonero, L., Clinton, K.M. & Johnson, A.G. (2011) On-road and simulated driving: Concurrent and discriminant validation. *Journal of safety research*, 42, 267-275.

McDonald, M. & Brackston, M. (1997) The role of the instrumented vehicle in the collection of data on driver behaviour. *IEE Colloquium on Monitoring of Driver and Vehicle Performance*, 122, 7/1-3.

Meijman, T.F. & O'Hanlon, J.F. (1984) Workload. An introduction to psychological theories and measurement methods IN: P.J.D. Drenth, H. Thierry, P.J. Willems & de Wolff, C.J., eds. *Handbook of Work and Organizational Psychology*, New York, USA: Wiley, 257-288.

Meister, D. (1976) *Behavioral foundations of system development*. New York: Wiley.

Michon, J.A. (1985) A critical view of driver behavior models: what do we know, what should we do? IN: L. Evans & R.C. Schwing, eds. *Human behavior & traffic safety*, New York: Plenum Press, 485-524.

Minin, L., Benedettob, S., Pedrottid, M., Rec, A. & Tesauria, F. (2011) Measuring the effects of visual demand on lateral deviation: A comparison among driver's performance indicators. *Applied Ergonomics*, in Press.

Miura, T. (1986) Coping with situational demand: A study of eye movements and peripheral vision performance IN: A. Gale, ed. *Vision in vehicles*, Amsterdam: North Holland, 205-217.

Moray, N., Dessouky, M.I., Kijowski, B.A. & Adapathya, R. (1991) Strategic behaviour, workload, and performance in task scheduling. *Human Factors*, 33 (6), 607-629.

Mourant, R.R. & Rockwell, T.H. (1970) Mapping eye-movement pattern to the visual scene in driving: An exploratory study. *Human Factors,* 12, 81-87.

Mulder, G. (1986) The concept and measurement of mental effort IN: G.R.J. Hockey, A.W.K. Gaillard & M.G.H. Coles, eds. *Energetics and human information processing*, Dordrecht, The Netherlands: Martinus Nijhoff Publishers, 175-198.

Myles, P.S. & Cui, J. (2007) Using the Bland-Altman method to measure agreement with repeated measures. *British Journal of A*, 99 (3), 309-311.

Nakayama, O., Futami, T., Nakamura, T. & R., B.E. (1999) Development of a steering entropy method for evaluating driver workload. IN: *International Congress and Exposition*, Detroit, Michigan, , March 1-4. Society of Automotive Engineers.

Neumann, D.L. & Lipp, O.V. (2002) *Spontaneous and reflexive eye activity measures of mental workload*, School of Psychology, The University of Queensland, Australia.

Norman, D.A. & Bobrow, D.G. (1975) On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44-64.

Nunes, L. & Recarte, M.A. (2002) Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5 (2), 133-144.

O'Donnell, R.D. & Eggemeier, F.T. (1986) Workload assessment mothodology IN: K.R. Boff, L. Kaufman & J.P. Thomas, eds. *Handbook of perception and human performance*, New York: Wiley, 42/41-42/49.

Olsson, S. & Burns, P.C. (2000) Measuring driver visual distraction with a peripheral detection task. IN: *NHTSA Driver Distraction Internet Forum*, Washington, DC.

Osborne, J. 2002, *Notes on the use of data transformations*, *Practical Assessment, Research & Evaluation*, vol. 8, no 6, pp. 1-8<http://PAREonline.net/getvn.asp?v=8&n=6>.

Osborne, J.W. & A., O. (2004) The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation*, 9 (6), 1-10.

Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H., Jamson, S., Mouta, S., Carvalhais, J., Santos, J., Anttila, V., Sandberg, H., Luoma, J., de Waard, D., Brookhuis, K., Johansson, E., Engström, J., Victor, T., Harbluk, J., Janssen, W. & Brouwer, R. (2004) *HASTE deliverable 2: HMI and safety-related driver performance*, Institute for Transportation Studies, University of Leeds.

Paelke, G.M. (1993) A comparison of route guidance destination entry methods. IN: *Human Factors and Ergonomics Society 37th Annual Meeting*, 1. 569-573.

Parasuraman, R. & Hancock, P.A. (2001a) Adaptive control of mental workload IN: P.A. Hancock & P.A. Desmond, eds. *Stress, workload, and fatigue*, London: Lawrence Erlbaum Associates, 305-320.

Parasuraman, R. & Hancock, P.A. (2001b) Adaptive control of mental workload IN: P.A. Hancock & P.A. Desmond, eds. *Stress, sorkload, and fatigue*, London: Lawrence Erlbaum Associates, 305-320.

Parkes, A. (1991a) Data capture techniques for RTI usability evaluation IN: Commission of the European Communities, ed. *Advanced telematics in road transport, Proceedings of the DRIVE Conference*, Amsterdam: Elsevier, 1440-1456.

Parkes, A. (1991b) The effects of hands-free telephone use on conversation structure and style. IN: *the 24th Annual Conference of the Human Factor Association of Canada*, Vancouver, BC. 141-146.

Pastor, G., Tejero, P., Choliz, M. & Roca, J. (2006) Rear-view mirror use, driver alertness and road type: An empirical study using EEG measures. *Transportation Research Part F: Traffic Psychology and Behaviour*, 9 (4), 286-297.

Pauzie, A. (2008) Evaluating driver mental workload using the Driving Activity Load Index (DALI). IN: *European Conference on Human Centered Design for Intelligent Transport Systems*, Loyn, France, 3th-4th April, 2008. 67-77.

Pauzie, A. & Pachiaudi, G. (1997) Subjective evaluation of the mental workload in the driving context IN: J.A. Rothengatter & E.C. Vaya, eds. *Traffic and transport psychology: Theory and application*, Amsterdam: Pergamon, 173-182.

Pitts, M.J., Williams, M.A., Wellings, T. & Attridge, A. (2009) Assessing subjective response to haptic feedback in automotive touchscreens. IN: *the First Interantional Conference on Automotive User Interfaces and Interactive Vehiclar Applications*, Essen, Germany. 11-18.

Pohlmann, S. & Traenkle, U. (1994) Orientation in road traffic. Age-related differences using an in-vehicle navigation system and a conventional map. *Accident analysis and prevention*, 26 (6), 689-702.

Posner, M.I. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.

Poysti, L., Rajalin, S. & Summala, H. (2005) Factors influencing the use of cellular (mobile) phone during driving and hazards while using it. *Accident Analysis & Prevention*, 37, 47-51.

Putrevu, S. (2001) Exploring the origins and information processing differences between men and women: Implications for advertisers. *Academy of Marketing Science Review*, 10.

Rakauskas, M.E., Gugerty, L.J. & Ward, N.J. (2004) Effects of naturalistic cell phone conversations on driving performance. *Journal of safety research*, 35 (4), 453-464.

Rakauskas, M.E., Ward, N.J., Boer, E.R., Bernat, E.M., Cadwallader, M. & Patrick, C.J. (2008) Combined effects of alcohol and distraction on driving performance. *Accident analysis and prevention*, 40, 1742-1749.

Ranney, T.A., Mazzae, E., Garrott, R. & Goodman, M.J. 2000, *NHTSA driver distraction research: Past, present and future*, NHTSA, <http://www-nrd.nhtsa.dot.gov/departments/Human%20Factors/driver-distraction/PDF/233.PDF>.

Ranney, T.A., Scott Baldwin, G. H., Vasko, S.M. & Mazzae, E.N. (2009) *Measuring Distraction on Potential of Operating In-Vehicle Devices*, National Highway Traffic Safety Administration.

Rantanen, E.M. & Goldberg, J.H. (1999) The effect of mental workload on the visual field size and shape. *Ergonomics*, 42 (6), 816-834.

Rasmussen, J. (1987) Cognitive control and human error mechanisms IN: J. Rasmussen, K. Duncan & J. Jeplat, eds. *New technology and human error*, Chicester, UK: John Wiley & Sons, 53-61.

Rayner, K. (1998) Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124 (3), 51.

Reasons, J., Manstead, A., Stradling, S., Baxter, J. & Campel, K. (1990) Errors and violations on the roads: a real distinction? *Ergonomics*, 33 (10/11), 1313-1332.

Recarte, M.A. & Nunes, L.M. (2000) Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied*, 6 (1), 31-43.

Recarte, M.A. & Nunes, L.M. (2003) Mental Workload While Driving: Effects on Visual Search, Discrimination, and Decision Making. *Journal of Experimental Psychology: Applied*, 9 (2), 119-137.

Recarte, M.A., Perez, E., Conchillo, A. & Nunes, L.M. (2008) Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish Journal of Psychology*, 11 (2), 374-385.

Reed, M.P. & Green, P.A. (1999) Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics*, 42 (8), 1015-1037.

Reger, M.A., Welsh, R.K., Watson, G.S., Cholerton, B., Baker, L.D. & Craft, S. (2004) The relationship between neuropsychological functioning and driving ability in dementia: A meta-analysis. *Neuropsychology*, 18 (1), 85-93.

Reid, G.B. & Nygren, T.E. (1988) The subjective workload assessment technique: a scaling procedure for measuring mental workload IN: P.A. Hancock & N. Meshkati, eds. *Human mental workload*, Amsterdam: Elsevier.

Riggs, L.A., Merton, P.A. & Morton, H.B. (1974) Suppression of visual phosphenene during saccadic eye movements. *Vision Research*, 14, 997-1010.

Rizzo, M. (2004) Safe and unsafe driving IN: M. Rizzo & P.J. Eslinger, eds. *Principles and practice of behavioral neurology and neuropsychology*, Philadelphia, USA: Saunders.

Rizzo, M., McGehee, D.V., Dawson, J.D. & Anderson, S.N. (2001) Simulated car crashes at intersections in drivers with Alzheimer disearse. *Alzheimer disease and associated disorders*, 15 (1), 10-20.

Rockwell, R.H. (1972) Eye movement analysis of visual information acquisition in driving. *Australian Road Research Board Proceedings*, 6 (3), 316-331.

Rockwell, T.H. (1988) Spare visual capacity in driving--Revisited. IN: A.G. Gale, M.H. Freeman, C.M. Haslegrave, P. Smith & S.P. Taylor, eds. *The second international conference on Vision in Vehicles*, Nottingham, England. Elsevier Science, North-Holland, 317-324.

Rogers, R.D. & Monsell, S. (1995) Costs of a predictible switch between simple cognitive tasks. *journal of Experimental Psychology: General*, 124 (2), 207-231.

Rognin, L., Alidra, S., Val, C. & Antoine, L. (2007) Occurrence of secondary tasks and quality of lane changes IN: D. Harris, ed. *Engineering Psychology and Cognitive Ergonomics*, Springer-Verlag Berlin Heidelberg, 397-406.

Rognin, L., Grimaud, I., Hoffman, E. & Zeghal, K. (2004) Assessing the impact of a new instruction on air traffic ontroller monitoring tasks. IN: *the International Conference on Human-Computer Interaction in Aeronautics*. AAAI Press, Toulouse, France.

Rosenbaum, G. (2000) *NOVA Online*, Available from: <http://www.pbs.org/wgbh/nova/everest/exposure/braintest.html>.

Roth, A.J. (1983) Robust trend tests derived and simulated: Analogs of the Welch and Brown-Forsythe Tests. *Journal of American Statistical Association*, 78 (384), 972-980.

Rouse, W.B., Edwards, S.L. & Hammer, J.M. (1993) Modelling the dynamics of mental workload and human performance in complex systems. *IEEE Transactions on systems, man, and cybernetics*, 23, 1662-1671.

Russell, D.W. (1990) The analysis of psychophysiological data: mutivariate approaches IN: J.T. Cacioppo & L.G. Tassinary, eds. *Principles of Psychophysiology*, Cambridge, UK: Cambridge University Press.

Salmon, E., der Linden, M. Van, Collette, F., Delfiore, G., Maquet, P., Degueldre, C., Luxen, A. & Franck, G. (1996) Regional brain activity during working memory tasks. *Brain*, 119, 1617-1625.

Santos, J., Merat, N., Mouta, S., Brookhuis, K. & De Waard., D. (2005) The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. *Transportation Research Part F*, 8, 135-146.

Schrauf, M., Sonnleitner, A., Simon, M. & kincses, W.E. (2011) EEG Alpha Spindles as Indicators for Prolonged Brake Reaction Time During Auditory Secondary Tasks in a Real Road Driving Study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 55 (1), 217-221.

Schumacher, E.H., Seymour, Travis L., Glass, Jennifer M., Fencsik, David E., Lauber, Erick J., Kieras, David E., Meyer, David E. (2001) Virtually Perfect Time Sharing in Dual-Task Performance: Uncorking the Central Cognitive Bottleneck. *Psychological Science*, 12 (2), 101-108.

Schvaneveldt, R.W. (1969) Effects of complexity in simultaneous reaction-time tasks. *Journal of Experimental Psychology*, 81, 289-296.

Schvaneveldt, R.W., Reid, G.B., Gomez, R.L. & Rice, S. (1998) Modeling mental workload. *Cognitive Technology*, 3, 19-31.

Seeing Machines (2009) *FaceLAB5 User Manual*. Canberra, Australia.

Senders, J.W., Kristofferson, A.B., Levision, W., Dietrich, C.W. & Ward, J.L. (1967a) The attentional demand of automobile driving. *Highway Research Record*, 195, 15-33.

Senders, J.W., Kristofferson, A.B., Levison, W., Dietrich, C.W. & Ward, J.L. (1967b) The attentional demand of automobile driving. *Highway Research Record*, 195, 15-33.

Serafin, C., Heers, R., Tschirhart, M. & Ullrich, C. (2007) User experience in the U.S. and Germany of in-vehicle touch screens with integrated haptic and auditory feedback. IN: *SAE World Congress*.

Shehab, M. (2007). *The effects of heavy goods vehicles (HGVs) on driver behaviour while car-following*, University of Southampton.

Shiffrin, R.M. & Schneider, W. (1977) Controlled and automatic human information processing: II. perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84 (2), 127-190.

Shinar, D. & Gurion, B. (2008) Looks are (almost) everything: where drivers look to get information. *Human Factors*, 50 (3), 380-384.

Shinoda, H., Hayhoe, M.M. & Shrivastava, A. (2001) What controls attention in natural environments? *Vision Research*, 41, 3535-3536.

Sivak, M. (1996) The information that drivers use: is it indeed 90% visual? *Perception*, 25 (9), 1081-1089.

Sodnik, J., Dicke, C., Tomazic, S. & Billinghurst, M. (2008) A user study of auditory versus visual interfaces for use while driving. *International Journal of Human-Computer Studies,* 66 (5), 318-332.

Steinman, R.M. (2003) Gaze control under natural conditions. IN: L.M. Chalupa & J.S. Werner, eds. *The Visual Neurosciences.* , Cambridge, Massachusetts, USA: MIT Press, 1339–1356.

Strayer, D.L., Drews, F.A. & Crouch, D.J. (2006) A comparison of the cell phone driver and the drunk driver. *Human Factors,* 48 (2), 381-391.

Strayer, D.L., Drews, F.A. & Johnston, W.A. (2003) Cell phone induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied*, 9, 23-52.

Summala, H. (1981) Driver/vehicle steering response latencies. *Human Factors*, 23, 683-692.

Summala, H. (1998) Forced peripheral vision driving paradigm: Evidence for the hypothesis that car drivers learn to keep in lane with peripheral vision. *Vision in Vehicles-VI*.

Summala, H., Nieminen, T. & Punto, M. (1996) Maintaining lane position with peripheral vision during in-vehicle tasks. *Human Factors*, 38 (3), 442-451.

Sun, W.S., Baker, R.S., Chuke, J.C., Rouholiman, B.R., Hasan, S.A., Gaza, W., Stava, M.W. & Porter, J.D. (1997) Age-related changes in human blinks. *Investigative Ophthalmology & Visual Science*, 38 (1), 92-99.

Teigen, K.H. (1994) Yerkes-Dodson: a law for all seasons. *Theory & Psychology*, 4, 525-547.

Tijerina, L. (1996) *Final report - Program executive summary: Heavy vehicle driver workload assessment*, National Highway Traffic Safety Administration.

Tornros, J.E. & Bolling, A. (2006) Mobile phone use - effects of conversation on mental workload and driving speed in rural and urban environments. *Transportation Research Part F: Traffic Psychology and Behaviour*, 9 (4), 298-306.

Tornros, J.E.B. & Bolling, A.K. (2005) Mobile phone use--Effects of handheld and handsfree phones on driving performance. *Accident Analysis & Prevention*, 37 (5), 902-909.

Treisman, A. (1988) Features and objects: The fourteenth annual Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, 40, 201-237.

Tsang, P.S. & Velazquez, V.L. (1996) Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, 39 (3), 358-381.

Tsimhoni, O. & Green, P. (2001) Visual demand of driving and the execution of display-intensive in-vehicle tasks. IN: *Human Factors and Ergonomics Society 45th Annual Meeting*. 1586-1590.

Tsimhoni, O. & Green, P. (2003) Time-sharing of a visual in-vehicle task while driving: the effect of four key constructs. IN: *the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Park City, Utah, Julay 21-24, 2003. 113-118.

Tsimhoni, O., Smith, D. & Green, P. (2004) Address entry while driving: speech recognition versus a touch-screen keyboard. *Human Factors*, 46 (4), 600-610.

Tonros, J.E.B. & Bolling, A.K. (2005) Mobile phone use--Effects of handheld and handsfree phones on driving performance. *Accident Analysis & Prevention*, 37 (5), 902-909.

Uang, S.-T. & Hwang, S.-L. (2003) Effects on driving behavior of congestion information and of scale of in-vehicle navigation systems. *Transportation Research Part C: Emerging Technologies*, 11 (6), 423-438.

Underwood, G., Crundall, D. & Chapman, P. (2002) Selective searching while driving: the role of experience in hazard detection and general surveillance. *Ergonomics*, 45 (1), 1-12.

Van Orden, K. F., Jung, T.-P. & Makeig, S. (2000) Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology*, 52, 221-240.

van Orden, K.F., Limbert, W., Makeig, S. & Tzyy-Ping, J. (2001) Eye activity correlates of workload during a visiospatial memory task. *Human Factors*, 43 (1), 111-121.

Velichkovsky, B.M., Dornhoefer, S.M., Kopf, M., Helmert, J. & Joos, M. (2002) Change detection and occlusion modes in road-traffic scenarios. *Transportation Research Part F-Traffic Psychology and Behaviour*, 5, 11.

Veltman, J.A. & Gaillard, A.W.K. (1998) Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 5, 656-669.

Vicente, K.J., Thornton, D.C. & Moray, N. (1987) Spectral analysis of sinus arrhythmia: a measure of mental effort. *Human Factors*, 29, 171-182.

Victor, T.W., Harbluk, J.L. & Engstrom, J.A. (2005) Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F-Traffic Psychology and Behaviour*, 8, 167-190.

Vogel, K. (2003) A comparison of headway and time to collision as safety indicators. *Accident Analysis & Prevention*, 35, 427-433.

Wainer, H. (1976) Robust statistics: A survey and some prescriptions. *Journal of Educational Statistics*, 1 (4), 285-312.

Walker, J., Alicandri, E., Sedney, C. & Roberts, K. (1991) In-vehicle navigation devices: effects on the safety of driving performance. IN: *Vehicle navigation and information system conference*, Warrendale, PA. Society of Automotive Engineers, 499-525.

Wang, J.S., Knipling, R.R. & Goodman, M.J. (1996) The role of driver inattention in crashes: New statistics from the 1995 Crashworthiness Data System. IN: *40th Annual Proceddings of the Association for the Advancement of Automotive Medicine*, Vancouver, British Columbia. 377-392.

Wann, J.P. & Swapp, D.K. (2001) Where do we look when we steer and does it matter? *Journal of Vision*, 1 (3), 185.

Waugh, J.D., Glumm, M.M., Kilduff, P.W., Tauson, R.A., Smyth, C.C. & Pillalamarri, R.S. (2000) Cognitive workload while driving and talking on a cellular phone or to a passenger. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 6-General Sessions, 276-279.

Welch, B.L. (1951) On the comparison of several mean values: an alternative approach. *Biometrika*, 33, 330-336.

West, R., French, D., Kemp, R. & Elander, J. (1993) Direct observation of driving, self reports of driver behaviour, and accident involvement. *Ergonomics*, 36 (5), 557-567.

Wewerinke, P. & Hogema, J. (2003) *Modelling driver behaviour in combined longitudinal and lateral control*, TNO Human Factors, Soesterberg.

White, C.W. (1976) Visual masking during pursuit eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 469-478.

Wickens, C.D. (1984a) *Engineering psychology and human performance (2nd ed.)*. New York: HarperCollins.

Wickens, C.D. (1984b) Processing resources in attention IN: R. Parasuraman & D.R. Davies, eds. *Varieties of attention*, New York: Academic Press, 63-102.

Wickens, C.D. (1987a) Attention IN: P.A. Hancock, ed. *Human factors psychology*, Amsterdam: Elsevier.

Wickens, C.D. (1987b) Information processing, decision making and cognition IN: G. Salvendy, ed. *Cognitive engineering in the design of human-computer interaction and expert systems*, Amsterdam: Elsevier.

Wickens, C.D. (1992) *Engineering psychology and human performance*. New York: HarperCollins.

Wickens, C.D., Sandry, D.L. & Vidulich, M. (1983) Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, 25 (2), 227-248.

Wierwille, W.W. (1993a) Demands on driver resources associated with introducing advanced technology into the vehicle. *Transportation Research Part C*, 1 (2), 133-142.

Wierwille, W.W. (1993b) Visual and manual demands of in-car controls and displays IN: B. Peacock & W. Karowowski, eds. *Automotive ergonomics*, London: Taylor & Francis, 299-320.

Wierwille, W.W. & Tijerina, L. (1996) An analysis of driving accident narratives as a means of determining problems caused by in-vehicle visual allocation and visual workload IN: A.G. Gale, ed. *Vision in Vehicles V*, Amsterdam: North-Holland.

Wikman, A.-S., Nieminen, T. & Summala, H. (1998) Driving experience and time sharing during in-car tasks on roads of different width. *Ergonomics*, 41, 358-372.

Wilson, G.F. & Eggemeier, F.T. (1991) Psychophysiological assessment of workload in multi-task environments IN: D.L. Damos, ed. *Multiple-task performance*, London: Taylor & Francis, 329-360.

Wilson, G.F., Fullenkamp, P. & Davis, I. (1994) Evoked potential, cardiac, blink and respiration measures of workload in air-to-ground missions. *Aviation, Space, and Environmental Medicine*, 65, 100-105.

Wittmann, M., Kiss, M., Gugg, P., Steffen, A., Fink, M., Poppel, E. & Kamiya, H. (2006) Effects of display position of a visual in-vehicle task on simulated driving. *Applied Ergonomics*, 37 (2), 187-199.

Wood, C., Gray, R., Young, J., Summers, J., Torkkola, K. & Massey, N. (2003) Inattentional blindness while driving. IN: *Driving Assessment 2003: Second International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Park City, Utah, USA, 21-24, July, 2003.

Wright, P. (1974) The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology*, 59 (5), 555-561.

Wynn, T. & Richardson, J.H. (2008) Comparison of subjective workload ratings and performance measures of a reference IVIS task. IN: *the Europe Conference on Human Centred Design for Intelligent Transport System*, Lyon, France, 3-4, April, 2008. 79-88.

Yang, Y., McDonald, M. & Zheng, P. (2009) New investigation to the driving performance affected by task designs and in-vehicle audio interfaces. *in revising*.

Yarbus, A. (1967) *Eye movement and vision*. New York: Plenum Press.

Young, M.S., Mahfoud, J.M., Stanton, N.A., Salmon, P.M., Jenkins, D.P. & Walker, G.H. (2009) Conflicts of interest: The implications of roadside advertising for driver attention. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12 (5), 381-388.

Young, M.S. & Stanton, N.A. (2004) Taking the load off: investigation of how adaptive cruise control affects mental workload. *Ergonomics*, 47 (9), 1014-1035.

Yu, C.-h. (2011) *Parametric tests*, Available from: <http://www.creative-wisdom.com/teaching/WBI/parametric_test.shtml>.

Zeghal, K., Grimaud, I., Hoffman, E. & Rognin, L. (2002) Delegation of spacing tasks from controllers to flight crew: Impact of controller monitoring tasks. IN: *the IEEE/AIAA Digital Avionics Systems Conference, 2B3*. IEEE Press, Irvine, CA, 1-9.

Zhang, H. (2004) The optimality of Naive Bayes. IN: *the 17th Florida Artificial Intelligence Research Society Conference*. the AAAI Press, 562-567.

Zheng, P. (2003). *A microscopic simulation model of merging operation at motorway on-ramps*, University of Southampton.

Zheng, P., McDonald, M. & Pickering, C. (2008) Effects of intuitive voice interfaces on driving and in-vehicle task performance. IN: *the 11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, China, 12-15, October. 610-615.

Zijlstra, F.R.H. (1993). *Efficiency in work behavior. A design approach for modern tools*, Delft University of Technology, the Netherlands.

Zimmerman, D.W. (1994) A note on the influence of outliers on parametric and nonparametric tests. *Journal of General Psychology*, 121 (4), 391-401.

Zimmerman, D.W. (1995) Increasing the power of nonparametric tests by detecting and downweighting outliers. *Journal of Experimental Education*, 64 (1), 71-78.

Zimmerman, D.W. (1998) Invalidation of parametric and nonparametric statistical tests by concurrent violation of two assumptions. *Journal of Experimental Education*, 67 (1), 55-68.

Zwahlen, H.T., Russ, A., Roth, J.M. & Schnell, T. (2003) Viewing Ground-Mounted Diagrammatic Guide Signs Before Entrance Ramps at Night: Driver Eye Scanning Behavior *Transportation Research Record*, 1843, 61-69.

Appendix A: Driver Questionnaire

Driver Questionnaire

You are invited to participate a study to investigate impacts of performing in-vehicle secondary tasks on drivers' visual behaviour and performance, in which you will be required to drive an instrumented vehicle in traffic (on urban roads and motorway), and to complete selected in-vehicle tasks (e.g. answering verbal puzzle questions or using a touch-screen). The aim of this questionnaire is to gather general information about your driving behaviour as part of recruiting process.

This questionnaire is divided into three sections. Section A specifies your personal information, Section B your driving experience, and Section C is about your experience of using in vehicle systems. Your answer will help us grouping drivers according to driving experience and behaviour. Please answer these questions as accurately as you can.

Section A About You

For the following questions, please tick ONE box only (tick the checkbox by clicking on it if you are filling this questionnaire electronically).

A1) Gender: (Please tick one box only)

Male ☐ Female ☐

A2) Age: (Please tick one box only)

25 - ☐	35-39 ☐	50-54 ☐	65 + ☐
25-29 ☐	40-44 ☐	55-59 ☐	
30-34 ☐	45-49 ☐	60-64 ☐	

A3)Do you drive with glasses or contact lenses (for the purpose of setting up the eye_tracking equipment): (Please tick one box only)

No ☐ Yes, with glasses ☐ Yes, with contact lenses ☐

Section B Driving experience

B1) Do you hold a full valid driving license use in the UK: (Please tick one box only)

Yes ☐ No ☐

B2) How many years have you been driving in the EU (including the UK)? (Please write in)

_______Years

B3) On average, what is your approximate annual mileage (× 1000)? (Please tick one box only)

5 ☐ 5 - 10 ☐ 10 - 15 ☐ 15 - 20 ☐ 20+ ☐

B4) On average, how many days do you drive per week? (Please tick one box only)

1 - ☐ 1 - 2 ☐ 3 - 4 ☐ 5 - 6 ☐ 7 ☐

Section C Experience of in-vehicle systems

C1) How often do you use hand-free mobile phone whilst driving? (Please tick one box only)

15+ min/day ☐ 6-15 min/day ☐ 0-5 min/day ☐ Sometimes ☐ Never ☐

C2) Do you consider answering hand-free mobile phone whilst driving to be distracting? (Please rate on a scale from 1 to 5)

Not distracting at all				Very distracting
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

C3) How often do you adjust car radio/CD or similar in-vehicle equipment while driving? (Please rate on a scale from 1 to 5)

Often				Never
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

C4) How safe do you consider the activities in D4 to be while driving? (Please rate on a scale from 1 to 5)

Safe				Unsafe
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

Please provide your contact details if you would like to participate the experiment. One of our researchers will contact you by email or telephone: (Please write in capital letters)

Name: __________

E-mail Address: __________

Address: __________

____

Post Code: __________

Contact Telephone Number: __________

Thank you for completing this questionnaire. Please send the completed questionnaire back to us using the freepost envelope provided.

Thank you again for your time and cooperation!

Data Protection Act 1998. The information you provide will only be used for research purposes. Any information you give will be treated in the strictest confidence and no information will be published that allows individuals to be identified. Your address will only be processed if you wish to volunteer to take part in further experimental trials. No organisations outside this project will have access to your personal data.

Appendix B: Dual-Task Activation Rating

Participant: __________________________

Data and Time: __________________________

The **Dual-Task Activation Rating** is to collect your evaluation on the workload (or demand) experienced for performing the secondary tasks. During the experiment you will perform different in-vehicle secondary tasks (e.g. operating on an in-vehicle system) whilst driving. You may feel some difficulties/demands while performing these tasks. Immediately after each secondary task performing, the experimenter will ask you to give a rating for your overall experienced workload/demand on this scale. You should give your rating by comparing your experienced workload for the current secondary task with the reference task - "***Turning on the Radio***", which has a fixed workload value of ***100***., i.e. the rating of 50 would mean the current task is as half demanding as the reference task; while a rating 200 would suggest it is twice as demanding as turning on the radio.

Note: You only have to give the rating vocally. An experimenter sitting in the rear seat will record each of your rating manually.

Task Type	Task	Rate
Reference	Tuning Radio	100
Auditory Task	Auditory Task Level 1_1st	
	Auditory Task Level 2_1st	
	Auditory Task Level 3_1st	
	Auditory Task Level 1_2nd	
	Auditory Task Level 2_2nd	
	Auditory Task Level 3_2nd	
	Auditory Task Level 1_3rd	
	Auditory Task Level 2_3rd	
	Auditory Task Level 2_3rd	
Visual Task	Visual Task Level 1_1st	
	Visual Task Level 2_1st	
	Visual Task Level 3_1st	
	Visual Task Level 1_2nd	
	Visual Task Level 2_2nd	
	Visual Task Level 3_2nd	
	Visual Task Level 1_3rd	
	Visual Task Level 2_3rd	
	Visual Task Level 2_3rd	

Appendix C: Post-Driving Interview

Thank you for finishing the driving session for the current study. This post-driving interview is aimed at gathering information on your opinion towards distraction in different driving scenarios, and awareness of your driving behaviour.

The interview includes three sections: Section A is about your experience of performing secondary tasks whilst driving, Sections B and C relate to detailed behaviour changes while performing these tasks. Your answer will be used to help us understand your individual driving experience and behaviour during the experiment.

Please answer these questions as accurately as you can.

SECTION A: EXPERIENCE OF SECONDARY TASKS

A1) Compared with normal driving, how much more demanding was the *auditory tasks* when *following another vehicle*? (Please rate on the scale 1-5, where 5 is most demanding)

No more demanding Most demanding

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

A2) Compared with normal driving, how much more demanding was the *auditory tasks* when *NOT following*? (Please rate on the scale 1-5, where 5 is most demanding)

No more demanding Most demanding

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

A3) Compared with normal driving, how much more demanding was the *visual tasks* when *following another vehicle*? (Please rate on the scale 1-5, where 5 is most demanding)

No more demanding Most demanding

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

A4) Compared with normal driving, how much more demanding was the *visual tasks* when *NOT following*? (Please rate on the scale 1-5, where 5 is most demanding)

No more demanding Most demanding

1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

SECTION B: YOUR DRIVING BEHAVIOUR WHEN PERFORMING **AUDITORY** TASKS

Were you aware of any of the following changes in your driving behaviour when performing the auditory tasks? (You can tick more than one)

1. Reduced speed	Yes	☐	Sometimes	☐	No	☐	Other	☐
2. Increase headway (distance to the vehicle ahead)	Yes	☐	Sometimes	☐	No	☐	Other	☐
3. Did you start to reduce speed and/or increase headway earlier (compared to when the tasks started)?	Yes	☐	Sometimes	☐	No	☐	Other	☐
4. Increase wheel movement adjustments	Yes	☐	Sometimes	☐	No	☐	Other	☐
5. Looked less at adjacent traffic	Yes	☐	Sometimes	☐	No	☐	Other	☐
6. Reacted more slowly	Yes	☐	Sometimes	☐	No	☐	Other	☐
7. Other (please describe)	☐							

Description:

SECTION C: YOUR DRIVING BEHAVIOUR WHEN PERFORMING **VISUAL** TASKS

Were you aware of any of the following changes in your driving behaviour when performing the visual tasks? (You can tick more than one)

1. Reduced speed	Yes	☐	Sometimes	☐	No	☐	Other	☐
2. Increase headway (distance to the vehicle ahead)	Yes	☐	Sometimes	☐	No	☐	Other	☐
3. Did you start to reduce speed and/or increase headway earlier (compared to when the tasks started)?	Yes	☐	Sometimes	☐	No	☐	Other	☐
4. Increase wheel movement adjustments	Yes	☐	Sometimes	☐	No	☐	Other	☐
5. Looked less at adjacent traffic	Yes	☐	Sometimes	☐	No	☐	Other	☐
6. Reacted more slowly	Yes	☐	Sometimes	☐	No	☐	Other	☐
7. Other (please describe)	☐							

Description:

Date:
Time:
Subject Number:

Appendix D: Consent form

UNIVERSITY OF Southampton
School of Civil Engineering and the Environment

Ethics Ref: CEE 2009/10-08
GRO Ref: 7302

Instrumented Vehicle Experiment

Consent form

The project aims to study drivers' visual and driving behaviour while performing in-vehicle secondary tasks, in order to study how drivers' visual and driving behaviour changed accordingly when performing secondary tasks. The findings from this research will benefit the advanced driver training and establishing a real-time, objective method of monitoring drivers' workload potentially. You have been asked to take part in a driving trial investigation. To show that you have agreed to take part in this trial we require you to complete this consent form.

SECTION A INFORMED CONSENT

I ………………………………………………… have received both verbal and written explanation of the study, and have also been given the opportunity to ask for clarification and/or further details should I wish.

I have freely given my consent to take part in this study. I understand my participation is voluntary and that I have the right to withdraw at any time and also for my data to be removed at a later date should I so wish. Finally I understand that all data will be used in accordance with the Data Protection Act 1998.

I understand that this trial will involve driving an Instrumented Vehicle on public roads, performing verbal and visual tasks which are no more onerous than the equivalent of tuning a radio, and a questionnaire survey to gather information about my experience of performing the secondary tasks in the trial.

I understand that it is my responsibility to drive in a manner that does not endanger other road users or myself.

Please tick the appropriate box:

☐ I confirm that:

- I am over the age of 25,
- I hold a full valid driving licence for use in the UK for more than 3 year,
- I have no current endorsements/penalty points on my driving licence,
- I have no medical condition or disability that precludes me from driving a motor vehicle on the public highway.

OR

☐ I do not fulfil all the above requirements but have been given dispensation to drive the University of Southampton instrumented vehicle *(this will be confirmed to you by the person administrating the test)*

I have read and understood the above information. By signing below and returning this form, I am consenting to participate in this study.

Signed: ……………………………………………. **Date:** ………………….

Print name: ……………………………………..

Continued overleaf…

SECTION B PARTICIPATION IN FURTURE TRIALS

B1) This project may be require participants for future driving trials in a similar test. Are you willing to be contacted in future to be asked to take part in further trials? (Please tick one box only)

Yes ☐ No ☐ End of questionnaire

B2) Please write in your preferred contact details so we may contact you regarding future trials. (Please write in)

Address: ....

....

Postcode: ....

Tel. Number: ....

Email: ....

Thank you for your time and cooperation

Data Protection Act 1998. The information you provide will only be used for research purposes. As with any data in this trial, any information you give as a volunteer will be treated in the strictest confidence and no information will be published that allows individuals to be identified. Your address will only be processed if you wish to volunteer to take part in further vehicle trials. This research is being carried out for the project by University of Southampton. No organisations outside of the project will have access to your personal data.

To be completed by person administrating the trial

I confirm I have seen this individual's driving licence and they fulfil the standard requirements to drive the instrumented vehicle

OR

I confirm I have seen this individual's driving licence and they have been given special dispensation to drive the instrumented vehicle because:

....

....

Those not meeting the requirements listed in Section A can be given special dispensation to drive the instrumented vehicle if:

- They are under 25 but over 21 and they have been cleared to drive by Head of School
- They have less than 3 years driving experience and have been cleared by Head of School
- They have endorsements up to three penalty points on their licence for minor offence. The following website can be consulted regarding motoring offences that carry penalty points.

http://www.direct.gov.uk/en/Motoring/DriverLicensing/EndorsementsAndDisqualifications/DG_10022425

If in any doubt check with the insurance office (Ruth McFadyen ext. 22417). Insurance office should be informed of all drivers who do NOT meet standard requirements.

Signed: …………………………………………. Date: ………………….

Appendix E: Driver Instructions

1. General Instructions

Thank you for participating in this experiment, which will help us to understand driving performance and visual behaviour while performing in-vehicle tasks when driving. You will receive instructions at the beginning of the experiment session. The tasks to be performed will be demonstrated before the driving session starts. Participation is expected to last approximately 2 hours and 30 minutes. ***Safety is paramount and must override any tests as indicated in Section 3 and 4.***

2. Instruction on Driving

The TRG Instrumented Vehicle you will use is a Fiat Stilo (automatic). You will be given a practice drive of about 30 minutes to become familiar and comfortable with the steering and dynamics of the Instrumented Vehicle. Then you will be asked to drive a pre-determined route, which will be detailed by the experimenter.

The practice drive is on motorway **M3,** during which you will be instructed to drive on the nearside lane, follow a passenger car ahead and try to keep the ***headway*** *(the distance between you and the vehicle ahead)* ***as close as you feel comfortable with, and as constant as possible.***

After you are familiar with the vehicle, the experiment will start with a route on the ***A34*** and ***A33***. ***On the A34***, you will be asked to follow a lorry in the inner lane at a ***headway as close as you are comfortable with, and try to keep it as constant as possible***. You need to follow the leading vehicle without changing lanes until the following process terminates. If the leading vehicle changes lane, the following process ends, and the experimenter may identify another vehicle to follow. While following, you will be instructed to perform the secondary tasks that have been demonstrated before. ***On the A33***, you will be instructed drive without following. You will be required to ***keep your speed as constant as possible at 55 miles per hour***, whilst a series of secondary tasks will be given to you.

You will be accompanied at all time by a safety observer sitting in the passenger seat and an experimenter in the rear seat. Please drive as you would normally do, when the following-session is over.

3. General Instruction on In-vehicle Secondary Tasks

You will be asked to perform some in-vehicle tasks. The tasks you are going to carry out in the experiment are listed below. The tasks will be demonstrated by the experimenter before you start driving.

Whilst driving, following the instruction from the experimenter sitting in the vehicle, you should perform these tasks only when you feel it is safe to do so. **You MUST abort any secondary task if at any time you feel unsafe**.

Immediately after each task, you will be asked questions about your experiences.

4. Instructions on Each Specific Secondary Task

There are 4 types of secondary tasks ***(at the testing stage)***:

1. ***A Simulated IVIS (In-Vehicle Information System) task.***
2. ***Sound Counting task;***
3. ***Visual Circle tasks;***

4. 1 In-vehicle Information System (IVIS) Task

There is one simulated in-vehicle information task: select a radio station using the touch-screen. Once the control interface is displayed on the LTC, you will need to turn on the radio and select "BBC 1". You need to click "Radio" button on the fist display, and then "FM Radio" and "BBC1" on the second one. A short radio clip will be played after you have correctly completed the procedure;

4.2 Sound Counting Task

In Sound Counting Task, you will be given one or more ***"Target Sounds"*** (about 320ms long each) at the beginning of each Sound Counting task, followed by 15 similar sounds appearing in at 2 second intervals. You need try to remember the "target sounds" and then count how many times each target sound has occurred in the following 15 sounds. You will be asked to report the counting results after each task.

There are three levels of Sound Counting Task: level 1 includes 1 target sound; In level 2, there are 2 targets, you need therefore count the times of each target appearing separately, and in level 3, 3 target sounds will be given. At the beginning of each Sound Counting task, there will be an auditory cue which can provide you the information of the level of the current task, e.g. "Sound counting task 2-1" stands for this is a task number 1 in level 2 (i.e. 2 target sounds). You can use this information to optimize your attention.

4.3 Visual Circle Task

In this task, several images of a group of circles, with one (or more) distinctly larger than the others will be displayed on a touch-screen mounted in the vehicle. The display is divided into several sections, which are labelled by numbers from 1 to 4/6/8. There will be an auditory cue every time when a new image appears. When you hear this cue, you will need to:

- find the large-sized circle/circles and either click on it/them, or
- verbally report which sections the bigger circle/circles is/are in. You only need to report the number and then click on the "next" button on the display when you finish with each image.

1 second after you finish with the current image, another image will appear with the same auditory cues. The same process will be repeated.

An example of Circle Reference Task is shown below:

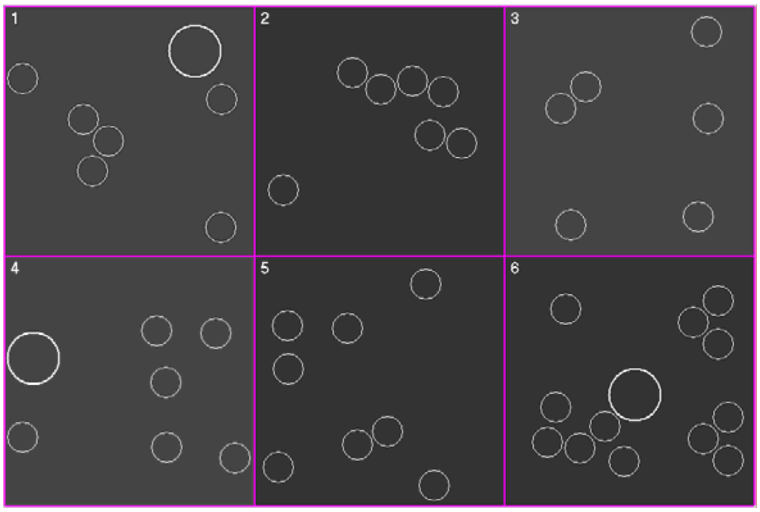


Figure 1: An example of Circle Reference Task

5. Route Instruction

The route instruction will be given by a safety observer who accompanies your driving at all time during the experiment.

Appendix F: Experimental Procedure

Name:	
Date:	
Time:	

Experimental Procedure

	Section	Task	Note
	Preparation	Preparation before a subject arrives	
1		Switch on IV power and data logger; Start logging	
2		Check for all IV sensors being properly functioning; stop logging	
3		Check recorded data and sufficient storage	
4		Switch on and run FaceLAB (including SceneView System)	
5		Check IR pod on, FaceLAB Camera connected, SceneView image quality, and sufficient storage; Send FaceLab data to SceneView laptop and test the data flow	
6		Connect and switch on Test Software laptop, test Operation Simulation System, check touch screen image and speakers	
7		Synchronise machine time for computers running IV logger, FaceLAB, SceneView System, and the Operation Simulation software to a time server (NTP)	
	Calibration	After the subject arrives	
8		Check Driving Licence and Fill in consent form	
9		Give driver instruction and demonstration	
10		Calibrate the Camera position, lock the position once adjusting finished (Instruction: Please look into the two cameras, make sure you can see reflections of your nose in both cameras in the centre. If not, we'll need to adjust the position of the cameras.)	
11		Set individual model for the driver and check tracking quality Instruction: Please look straight ahead, keep neutral expression and stable gaze	
12		Calibrate gaze accuracy (Instruction: Same as the previous procedure, look at left and right camera, keep neutral expression and keep your gaze possible.)	
13		Adjust World Model (Instruction: Please look at left, rear, right mirrors, speedometer, tachometer, touch screen and corners of the windscreen.)	
14		Calibrate gaze angle by 11 Points and record starting and ending time Calibration points are: Straight ahead, Left Camera, Right Camera, Left Mirror, Right Mirror, Rear Mirror, Touch Screen (4 corners and the central area)	
15		Calibrate driver's saccade and blink, and record starting and ending time Instruction: Please look at following point in turns	
16		Measure and record the position of the driver's head centre relative to the FaceLAB cameras	

	Test Run	Collect data on the two driving Scenarios	
17		Driver familiarisation session from the University to Winchester (about 30 mins)	
18		Perform reference task on the M3 at the end of familiarisation session Instruction: Please follow A CAR in front in the inner lane, try to keep the headway as constant as possible when performing the task	
19		Enquire the driver does he or she need longer familiarisation session, if yes, carry on driving on M3; otherwise, go to test road either A34 or A33 (order counterbalanced between subjects)	
20		On A34Instruction: Please follow A LORRY in front in the inner lane, try to keep the headway as constant as possible.	
21		on A33 Instruction: Free driving, keep the speed as constant as possible	
22		Performing all secondary tasks on each test route respectively, record subjective rating, answers for auditory tasks, and collect baseline data	
	Interview	Back to the University and Interview	
23		Park the IV and Interview the subject	
24		Pay the participation and sign the forms	
	Post-Test Run	Download data	
25		Switch off all equipment, charge the IV battery	
26		Backup Data at the first possible time	