Non-recurrent traffic situations and traffic information

Determining preferences and effects on route choice

Dissertation committee:

prof. dr. F. Eising	University of Twente, chairman/secretary
prof. dr. ir. B. van Arem	University of Twente, promotor
prof. dr. ir. E. C. van Berkum	University of Twente
prof. dr. ir. M. F. A. M. van Maarseveen	University of Twente, ITC
prof. dr. ir. S. P. Hoogendoorn	Delft University of Technology
prof. G. Lyons	University of the West of England
prof. dr. G. C. de Jong	University of Leeds
dr. M. van Noort	TNO
prof. dr. G. P. M. R. Dewulf	University of Twente (reserve)

TRAIL Thesis Series T2011/16, The Netherlands TRAIL Research School

TRAIL Research School
P.O. Box 5017
2600 GA Delft
the Netherlands
T: +31 (0) 15 278 6046
F: +31 (0) 15 278 4333
E: info@rsTRAIL.nl

CTIT Dissertation Series No. 11-203
Centre for Telematics and Information Technology
P.O. Box 217 - 7500 AE Enschede - the Netherlands

This thesis is the result of a Ph.D. study carried out between 2003 and 2011 at the University of Twente, faculty of Engineering Technology, department of Civil Engineering, Research Center Applications of Integrated Driver Assistance.

ISBN 978-90-5584-151-6
ISSN 1381-3617

Typeset in LaTeX

Printed by Haveka, Alblasserdam, the Netherlands.

Non-recurrent traffic situations and traffic information

Determining preferences and effects on route choice

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de Rector Magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op dinsdag 6 december 2011 om 14:45 uur

door

Thijs Johan Muizelaar

geboren op 26 januari 1980
te Emmen

Dit proefschrift is goedgekeurd door de promotor:

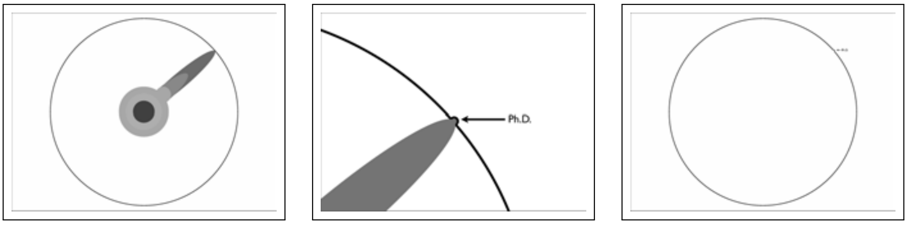
prof. dr. ir. B. van Arem

Voorwoord

Als me vooraf was gevraagd wat het lastigste zou zijn aan promoveren had ik waarschijnlijk geantwoord dat de inhoud en het doen van onderzoek de grootste uitdagingen zou zijn. Inmiddels, net iets meer dan 8 jaar verder, is mijn antwoord van een hele andere orde. Er voor zorgen dat je al die tijd met je aandacht bij dat ene onderwerp blijft en gemotiveerd te blijven om bij elke (kleine of grote) tegenslag door te gaan, daarin schuilt denk ik de ware uitdaging voor de promovendus. Dat dit proefschrift er ligt is voor mezelf dan ook vooral een bewijs dat ik daarin veel geleerd heb (naast ontelbaar veel andere dingen).

Een van de dingen die me daarbij heeft geholpen is de “Illustrated Guide to a PhD” van Matt Might (http://matt.might.net/articles/phd-school-in-pictures/) (figure 1). Stel je voor dat een grote cirkel alle kennis beschrijft, waarover je via je opleiding telkens een stukje leert. Uiteindelijk, na het behalen van je Master titel, heb je een specialisme opgebouwd en aardig wat kennis op dat terrein vergaard. Tijdens je promotie onderzoek bouw je daarop voort, door veel te lezen en te herhalen wat al is beschreven. Zo bereik je de grens van het kennisgebied van je eigen onderzoeksterrein (figure 1a). Dan volgt de uitdaging, het passeren van deze grens in onbekend gebied. Hier moet je het helemaal zelf doen. Uiteindelijk, na vele uren werken en doorzetten, doorbreek je die grens en creëer je een nieuw stukje kennis (figure 1b). Daarbij heb je je zo gefocust dat de wereld bestaat uit dat ene kleine stukje kennis. Die bijdrage is van groot belang, maar vergeet daarbij vooral niet om te kijken naar het grotere geheel (figure 1c).

Deze simpele maar geweldige plaatjes hebben me laten beseffen dat promoveren niet eenvoudig is en dat je je bewust moet zijn waar je focus ligt. Als je met je onderzoek bezig bent, moet je niet bij elke stap teruggrijpen naar het grotere geheel. Dan zie je vooral hoe klein je bijdrage aan de wetenschap en samenleving is. Op andere momenten is het juist goed om naar de context te kijken, om te beseffen dat je eigen onderzoek onderdeel is van de grote kenniscirkel. Gedurende de afgelopen jaren heb ik veel geleerd over verkeer, verkeersinformatie, routekeuze, het modelleren van gedrag en over alles wat komt kijken bij het doen van een promotie onderzoek. Daarbij hebben velen mij geholpen en een bijdrage geleverd.



(a) (b) (c)

Figure 1: *Selection of “The Illustrated Guide to a PhD (Might, 2011)*

Als eerste wil ik Bart van Arem, mijn promotor, bedanken. Mijn eerste kennismaking met Bart was tijdens mijn sollicitatiegesprek voor het doen van een promotie onderzoek bij AIDA. Ondanks mijn ietwat naïeve voorbereiding en opvallende kledij (zo begreep ik veel later) heb je me het vertrouwen gegeven voor deze plek. Dat vertrouwen heb ik daarna tijdens de uitvoering aardig op de proef gesteld, maar je hebt me altijd gesteund en geholpen als ik er niet uitkwam. Daarin zijn twee momenten erg belangrijk geweest. De eerste dat je me liet inzien dat mijn promotie onderzoek wel degelijk vernieuwende aspecten in zich heeft en het tweede over m'n proefschrift zelf. Na dat gesprek begrijp ik wat "Kill your darlings" echt betekent.

De internet enquête in mijn onderzoek is alleen mogelijk geweest met de hulp van Edouard Bunning van RM Interactive, de leden van de gebruikersgroep van kenniscentrum AIDA, JR Online en alle mensen die aan de enquête hebben meegewerkt. Ook wil ik de gebruikersgroep en ontwikkelaar van ModSurvey bedanken voor de ondersteuning bij het aanpassen van de software voor het maken van de enquête.

Voor het doen van het onderzoek naar het routekeuze gedrag was een simulator onontbeerlijk. Hiervoor ben ik grote dank verschuldigd aan Mark Raadsen en Albert Schoute als zijn afstudeerdocent. Mark heeft bij zijn afstuderen voor de opleiding informatica het grootste gedeelte van de AIDA RCS ontwikkeld en getest. Het was leuk om samen met jou na te denken over het wat en het hoe, en te kijken hoe we om moesten gaan met alle problemen met Paramics. Goed om te horen dat je inmiddels als software architect in het verkeersdomein bezig bent. Het experiment zelf heeft daarna even op zich laten wachten. Zonder alle deelnemers had ik geen resultaten gehad om te onderzoeken. Ik wil daarvoor alle deelnemers van de UT, TNO, TomTom, OMFL en vrienden en familie bedanken.

Tijdens mijn onderzoek is er ook buiten de UT interesse geweest en ben ik een aantal mensen dank verschuldigd. Zo wil ik van de collega promovendi een aantal mensen bedanken, in het bijzonder Geertje Hegeman en Maura Houtenbos. Eindelijk is Charlie klaar met z'n onderzoek. Casper Chorus wil ik ook bedanken voor het op weg helpen met de soms weerbarstige toepassingen van discrete keuze modellen. TNO en TomTom wil ik bedanken, in het bijzonder Rob van den Berg, Rob Schuurbiers, Freek Faber en Martijn van Noort. Ik heb in het begin en eind veel gehad aan alle feedback. Martijn, ik wil jou bijzonder bedanken voor de vele vrije uren die je hebt besteed aan het doorlezen van mijn concept proefschrift. Ik vind het een eer dat je in mijn promotie commissie zit. Namens TRAIL heeft Conchita me erg goed geholpen met het contact met de drukker en het controleren van het manuscript.

In de tijd dat ik in Enschede heb gezeten heb ik altijd gewerkt in een prettige werkomgeving, waarvoor alle (oud-)collega's verantwoordelijk zijn. Iedereen heel erg bedankt voor de fijne tijd in Twente: Eric, Martin, Mark, Frans, Wendy, Gio, Bart, Mako, Kasper, Bas, Tom, Martijn, Mohamed, Nina, Jing, Luc, Erwin en alle anderen. Cornelie, jou wil ik speciaal bedanken voor de geweldige kamergenoot die je was. Zoals je zelf al schreef, iemand met dezelfde voorkeur voor muziek en chocolade(eitjes). De gastvrijheid van jou en Steffen in Basel was echt super. Dorette wil ik bedanken voor alle hulp bij de administratieve zaken. Ook de afstudeerders Jeroen, Kubilay en Tim wil ik bedanken voor de fijne samenwerking.

Nadat ik bij de vakgroep verkeer ben weggegaan, ben ik aan de slag gegaan bij Capgemini. Ik heb binnen de vakgroep mobiliteit al veel interessante klussen gedaan. Ook de betrokkenheid en support bij het afronden van mijn onderzoek waardeer ik. Martijn, Jacco, Erik, Gio, Daan, Mark, Hilde en alle anderen, bedankt. Tijdens mijn opdracht bij Rijkswaterstaat in Lelystad heb ik een aantal mensen ontmoet die ik dankbaar ben voor de kansen die ze me hebben geboden en de fijne samenwerking. Jos, Margriet, Wilfried, Rudi en Bert, allemaal bedankt!

Tijdens een promotie onderzoek zijn vrienden en familie een grote hulp om te zorgen dat je af en toe met andere dingen bezig bent dan met je onderzoek. Zeker de afgelopen tijd is het er door het schrijven van het boekje bij ingeschoten om jullie vaak op te zoeken of hebben jullie wel eens een niet-zo-vrolijke Thijs meegemaakt. Jullie hebben me desondanks altijd een luisterend oor geboden, of je best gedaan om zorgvuldig het p-woord te vermijden. In het bijzonder Inge, Rolf, Manus, Dorine, Jan, Femke, Tamara en Stephan, bedankt! Ook in de kop van Noord-Holland hebben Tim, Els, Sandra, Paul, Linda en Rick altijd interesse getoond. Bedankt hiervoor, en dat rondje Afsluitdijk moet er nu echt van komen!

Een verdediging kan niet zonder paranimfen. Jan en Wietse, ik ben blij en trots dat jullie me terzijde staan tijdens de verdediging. Tijdens mijn onderzoek ben ik niet altijd de gezelligste en vrolijkste zoon en broer geweest, als ik er al was. Jullie waren altijd bereid om naar mijn verhalen, gezeur of onzin te luisteren. Dat ik ondertussen niet zoveel aandacht had voor wat jullie bezig houdt, hebben jullie nooit laten merken. Jan & Riet, ik ben er supertrots op dat ik jullie zoon ben. Jullie hebben me altijd gesteund in mijn keuzes en hebben me net zo eigenwijs gehouden als dat goed voor me is. Zonder jullie zou ik nooit zijn geweest wie ik nu ben.

Lieve Marion, samen met mij heb je het hele traject meegemaakt, met alle ups en downs die er zijn geweest. Ik heb het je niet makkelijk gemaakt de afgelopen jaren. De eindstreep was lang ver weg, maar nu is het eindelijk zo ver. We hebben bewezen dat we samen de hele wereld aankunnen. Ik kijk uit naar de komende jaren waarin we weer veel meer tijd samen hebben, op de racefiets, het trekvogelpad, op reis of in huis. Dankjewel voor alles.

Hilversum, november 2011

Thijs Muizelaar

Contents

List of Tables

List of Figures

The secret of getting ahead is getting started. The secret of getting started is breaking your complex overwhelming tasks into small manageable tasks, and then starting on the first one.

— Mark Twain

Chapter 1

Introduction

An often heard complaint in the Netherlands about traffic information on the radio: "the traffic jam I am in, is not mentioned at all!". Even though the introduction of personal navigation devices, which are able to receive information on traffic jams, meant a better informed driver, the current situation could benefit from further improvements with respect to traffic information. Particularly addressing the previous complaint, by personalising the traffic information would bring a benefit to a driver. In this thesis we focus on traffic information in non-recurrent traffic situations, for which we use the trip context and the personal characteristics. Before introducing the research in more detail, the chapter starts with a background on different traffic situations, reliability and uncertainty and traffic information.

1.1 Background

Land travel has changed dramatically over the course of centuries. Whereas in the times of the Romans lots of travel was done on foot and using cart and horses, the invention of the steam engine created possibilities for train travel. The innovation of the combustion engine marked the starting point for more individual forms of travel. Together with mass production, many people these days own at least one car. In the Netherlands, this results in a total of around 7.7 million passenger cars and 2.1 million commercial vehicles in 2011 (CBS, 2011) which is almost the double of the amount of vehicles in 1985 (Kennisinstituut voor Mobiliteitsbeleid, 2008). Together, these vehicles travel around 130 billion km/year in the Netherlands alone on a total road length of 140.000 km. On average each vehicle travels around 12.400 km/year.

As the car has become the favorite way of commuting, lots of vehicles (and people) are on the road during morning and evening rush hours. Increasingly, the road network is not capable of providing sufficient capacity for such an amount of vehicles. Congestion and delay are often the results. In the last 25 years, the use of cars in terms of vehicle kilometres has seen an increase of 54%, which is be explained by (Kennisinstituut voor Mobiliteitsbeleid, 2010):

1. an increase of the population size, which means more people can travel;
2. an increase of the travel distance, for example because people are living further away from their work location, and;
3. an increase of the frequency with which people travel, which is likely to be caused by more people increasingly joining activities away from home.

As the modal split between public transport and cars has not seen a change over the last years (with 75% of all travelled kilometers and around 50% of all trips made by car), it is easy to imagine the problems caused by the use of cars. Over the last years, the amount of congestion increased, which has not been compensated for by additions to the road network. In 2000, the total amount of lost vehicle hours was 44 million on the highways, which increased to 62,5 million in 2010, an increase of 42% (Ministerie van Verkeer en Waterstaat, 2010), which is around 8.6% of the total travel time on the highways. The main part of these lost vehicle hours occur during the rush hours when many people are commuting. Weekends show a smaller increase in amount of lost vehicle hours. The increase also varies between specific locations or areas. For example, certain highways in strongly urbanized areas such as the Randstad have a much larger increase than areas outside of the Randstad. Travel time has increased with around 2% and the standard deviation is increased with 4%. Both vary with time of day and month of the year. In the rush hours, the largest increases of the standard deviation of travel time are found.

Congestion leads to significant societal and economic costs, both directly and indirectly (Koopmans & Kroes, 2004). Direct costs are related to the congestion itself, as an hour of travel can be put to a monetary value (often called value of time). This value of time depends on the motive of travel, as a recreational traveller will often have a lower value of time compared to a traveller with a commute motive, or a freight operator which has to deliver its goods on time. Indirect costs are caused by the congestion, but not the congestion itself. For example, companies maintaining a larger stock because of unreliable travel times, or commuters leaving early to prevent experiencing delays caused by congestion.

Also other effects are of importance, such as effects on the environment, safety, etc. In 2009, the estimated total direct costs (Kennisinstituut voor Mobiliteitsbeleid, 2010) are between €2.4 and €3.2 billion for motorways alone. Reducing congestion costs has an immediate societal effect, as the productivity will increase. Reducing congestion and improving reliability of the transport system is thus an important target. This is also shown in the title of the Dutch government policy for traffic and transport: "Towards a reliable and predictable accessibility" (Ministerie van Verkeer en Waterstaat, 2005) which is further elaborated in the policy regarding optimizing the utilization of the existing network capacity (Ministerie van Verkeer en Waterstaat, 2008). However, building new infrastructure is not the only remedy to congestion. Improving the use of the currently available infrastructure is likely to cost less and be faster in achieving results. The ongoing development of information and communication technology is a main driver for the innovations in intelligent transport systems (ITS) and various subsystems such as advanced traveller information systems (ATIS). One of the goals of these systems is to aid drivers and road operators by increasing reliability and reducing uncertainty.

1.1.1 Reliability and uncertainty

In order to further explain the concept of reliability and uncertainty, we use a simplified representation of the transport system in figure 1.1. The transport system consists of travel demand and supply of infrastructure, which results in traffic flows on the infrastructure. If the travel demand exceeds the available supply of infrastructure (at a specific time and place in the transport system), it will result in congestion. This congestion again influences supply and demand, which over time will result in a “stable” situation in which supply and demand are balanced (possibly with congestion). However, in reality, the transport network is influenced by disturbances, either from within the system (endogenous) or from outside the system (exogenous) (Emmerink et al., 1995). These shocks are unpredictable events, which affect the “stable” situation. Endogenous shocks occur from within the system and have a relation to the demand and supply. For example, in a congested situation, the probability of an accident is higher. Exogenous shocks from outside the transport system directly affect the traffic flows, for example heavy rain. In turn, this can change the supply or demand, which also applies to endogenous shocks. These disturbance are often used to classify the cause of congestion as non-recurrent (as in Hall (1993); Hallenbeck et al. (2003); Enrique Fernández L. et al. (2009)).

Regardless of the source of a shock, the “stable” traffic flow is different when a disturbance occurs. This is likely to result in a trip which is not according to the expectations of a traveller. If that happens, a trip is said to be unreliable (Hilbers et al., 2004). Reliability in a transport system is described by the chance a trip can be made within a specific bandwidth of the expectations of the traveller for this trip, of which travel time is the most used expectation. Reliability can be separated into variation and predictability (Bates et al., 2001). Variation concerns the amount with which travel time varies over time or locations. Predictability is the level to which a traveller can predict a specific occurrence of travel time as a result of the current travel flows. If the variation is large and the predictability low, a trip is considered to be unreliable.

Both the level of variation and predictability are influenced by the behaviour of road operators and travellers. The degree of variation and predictability depends on the individual traveller. For example, a daily occurring traffic jam on the highway is predictable to a certain level and shows little variation to a commuter. The same traffic jam can mean a big variation in travel time and a highly uncertain arrival time for another traveller, who is not familiar with this traffic jam. Previous experiences and knowledge of the road network and the congestion that occurs are very important in dealing with (un)reliability. This introduces a new element in the previously mention concept of non-recurrent congestion; the context.

A traffic situation is called non-recurrent if it is irregular, unexpected or unknown; it has not occurred before in the current context. This context depends on the viewpoint (for example that of the road operator or traveller). In figure 1.1, this context is not displayed. Classifying a traffic situation to be recurrent or non-recurrent needs extra information besides the elements displayed in this figure. For example, a traffic situation which has been experienced before by a driver might

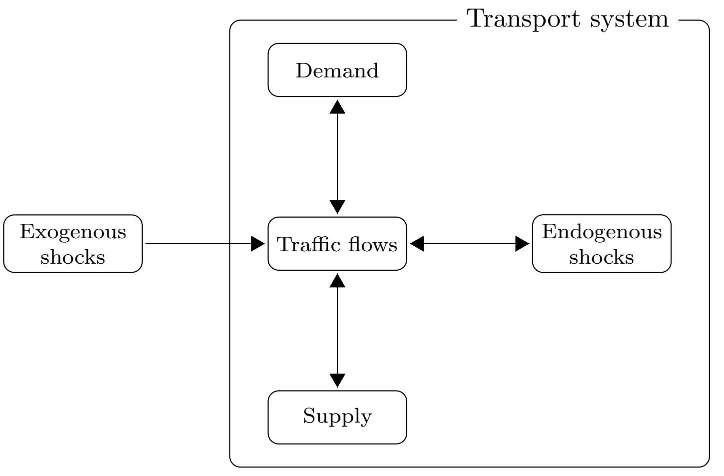


Figure 1.1: *Transport system with shocks (Emmerink et al., 1995)*

be recurrent, but another driver without that previous experience or knowledge, will call the same situation non-recurrent. No approach is currently available to include the context of the individual in describing traffic situations as recurrent or non-recurrent. However, Levinson (2003) already notes that especially during non-recurrent congestion the effect of traffic information on travel time savings is the largest. In this thesis, the focus is therefor on the non-recurrent traffic situations, in particular from the viewpoint of the driver.

Drivers can deal with (un)reliability in an individual trip by changing their behaviour and choices. A traveller might consider other destinations, leave early or late, or change the mode of travel. Travellers can also make decisions for their route, either before the trip, by minimising the likeliness of congestion, or en-route, to avoid (unexpected) congestion. In all situations, the level of non-recurrence affects these choices, together with personal characteristics. Next to these, also traffic information on the current (and expected) situations are taken into account when making a choice for the current trip. Information is expected to allow travellers to improve their choice making (Ministerie van Verkeer en Waterstaat, 2005). Advanced Traveller Information Systems aim at informing travellers on the current and future state of the transport system, to enable them to improve their travel experience.

1.1.2 Advanced Traveller Information Systems

Advanced Traveller Information Systems (ATIS) serves multiple goals, depending on the actor involved. For the government or road operator, ATIS contributes to (re)directing traffic in such a way that a more efficient use of the road network is achieved. This effect of ATIS might be achieved because the travellers on the network, and those planning to make a trip, are better informed about the status of the travel network. Travellers thus are expected to be able to make more conscious choices. Companies are also interested in ATIS, as it is a service or product that can be offered to travellers. It is in the interest of the company that travellers use their services or products with ATIS. ATIS can thus be used to make profit. Travellers use ATIS to improve their own travelling experience, by acquiring more personally efficient routes. On top of that, ATIS is also used to increase the comfort level during the trip, together with reducing, for example, travel distance or travel time (Adler & Blue, 1998). In general, drivers (and travellers) use traffic information to reduce uncertainty (Lappin, 2000), and more specifically to:

1. assess traffic congestion on their route;
2. judge the effects of incidents on their trip;
3. decide among alternate routes;
4. estimate their trip duration;
5. time their trip departure.

This list concerns a selected number of reasons related to a trip that is already planned to be undertaken by car. Broadening the scope should also lead to reasons relating to other choices, such as modality, actually making the trip and between various destinations or sequences hereof.

In recent years, many developments have taken place regarding ATIS. One of the first systems was tested in Berlin, at the end of 1989 (Sparmann, 1991). This system provided drivers with dynamic route guidance, based on historic floating car data, and measurements from traffic lights. Several other systems like this were tested, for example in the Netherlands with the RIC-project (Katteler & Broeders, 2002). This project aimed at making drivers familiar with the concept of in-car traffic information systems, which offered the drivers the ability to acquire current traffic information. 700 drivers had the opportunity to use such a traffic information system during the project, which started in 2000 and lasted one year. The information system used RDS-TMC (Radio Data System - Traffic Message Channel) to provide the actual traffic information and was able to display the information graphically (using a map) or textually.

Since the early nineties, many other developments regarding traffic information have taken place. Numerous Variable Message Sign (VMS) have been installed, especially on highways but also on the secondary road network. This VMS system informs drivers about current or expected travel times towards certain locations, or the current length of congestion on certain highways. This enables drivers to update their route choice. The Internet as a source of traffic information became available. Several websites in the Netherlands provide a realtime view of the status of the highways, such as the website of the ANWB. In the most recent years, an enormous growth of in-car navigation systems can be seen and a reduction of the costs of these systems and even "free" systems are available (on a smartphone). Most of these systems are able to inform the driver of congestion, and update the route if necessary. More recent developments in navigation devices introduced connected systems, which are continuously updated with traffic information using GPRS or UMTS. Quite often, this also allows information to be sent the other way, which means information about the vehicle's speed and location is uploaded to a service provider to be used in traffic information.

A large variety of systems is available and work to improve or extend the functionality is undertaken. Many improvements are possible, which is also mentioned by drivers (Ministerie van Verkeer en Waterstaat, 2004). The top three of elements of traffic information which can be improved are:

1. the information is not current;
2. the information does not mention short or daily congestion;
3. the information is too general and does not apply to personal trips.

However, drivers do not seem to be dissatisfied with the information provided, since the average value given to traffic information is a 7 on a scale of 1 to 10 (Ministerie van Verkeer en Waterstaat, 2004). This study indicates that one of the areas where improvement is possible is the personalization of ATIS. Personalized traffic information is based on the needs and preferences of the individual driver and can deal with varying circumstances.

ATIS has also been subject of interest by many scientists, as ATIS concerns many subjects in the field of traffic and transport research and outside. For example the presentation of information is important (Dicke et al., 2004; Dicke & Brookhuis, 2005; Brookhuis et al., 2008) and especially in relation to navigation systems and wayfinding (Burnett, 1998). Many other studies focus on the relation between choice behaviour and ATIS, especially in route choice behaviour.

Route choice is about making a choice for a sequence of roads that lead from an origin to a destination. The amount of possibilities for such a sequence is enormous but finite when excluding loops. The behavioural aspects involved in making a choice range from habit, learning behaviour, risk aversion to level of education and age. This means modelling route choice behaviour, in combination with traffic information is complex. As a consequence, much research has been dedicated to specific elements of route choice in combination with traffic information, for example learning (Chen & Jovanis, 2003; Avineri & Prashker, 2005; Bogers, 2009; Ben-Elia & Shiftan, 2010) and risk attitude (Bonsall, 2004b; Katsikopoulos et al., 2002; Palma & Picard, 2005). Bogers (2009) for example shows that the most recent experiences make up for 20% of the perceived travel time of a route. The attitude towards risk is of influence on the choices made. Some drivers are more willing to take a route with a larger uncertainty in the travel time then others. Bonsall (2004b) argues that depending on the attitude a traveller will apply different strategies to deal with this uncertainty.

Much of the research on route choice and ATIS has been dedicated to the use of traffic information as means of improving the performance of the road network. As such traffic information is used to “seduce” travellers to make those choices that will contribute to a better performance of the road network. On the other hand, the effect of traffic information on (route) choice behaviour will be largest if the information is personalized and is used as a means to improve the experience (not necessarily in terms of travel time) of the individual traveller. This involves not only taking into account the preferences of a traveller, but also the context of the current trip.

For instance, a traveller might not be interested in the current total travel time on certain alternatives on the highway network (as often displayed on variable message signs). These alternatives often do not completely apply to his trip as his destination may be halfway. If the congestion is at the latter part of the alternative next to his destination, the driver might be tempted to use another route, which would not be necessary. In short, if the current traffic situation is different from the expected or known situation (which we call non-recurrent, as recurring situations are likely to be known or expected), a traveller has the largest need for traffic information to enable making “better” choices. This does not just apply to highways, but also to the urban road networks where the possibility of using alternative routes is much larger and less likely to be communicated. Little is known about the preferences for such traffic information and for which situations this applies. Neither is it known how drivers will use such personalized traffic information and what the effects will be in these non-recurrent situations.

This thesis focuses on the combination of non-recurrent traffic situations and traffic information. In the remainder of this chapter, the specific research objective and research questions (section 1.2) are presented, followed by the scope (section 1.2.4) and the relevance of this research for science and society (section 1.3). The chapter ends with the research approach and an outline of this thesis (section 1.4).

1.2 Research objective and scope

Based on the background given in section 1.1, this thesis has the following main objective: ***To gain more insight into the impact of traffic information on route choice behaviour in non-recurrent traffic situations.***

It is expected that non-recurrent traffic situations have an influence on the preferences for traffic information in these situations, and that non-recurrent traffic situations influence the route choice behaviour in the presence of this traffic information. The objective consists of three main elements:

1. non-recurrent traffic situations;
2. traffic information, and;
3. route choice behaviour.

These three main elements are described in further detail. Each element will be described as part of the main objective using main and sub research questions.

1.2.1 Non-recurrent traffic situations

Non-recurrent traffic situations are traffic situations which have no similar or equal traffic situation in the past, where the history is specific to each individual. The individual can be both the individual traveller or a road operator, whereas the first is expected to regard the traffic situation for his trip, and the road operator is expected to regard the traffic situation for the whole road network. What might be recurrent for one traveller, might be a one-time only event for another. This justifies the importance of the individual in defining non-recurrent traffic situations.

Traffic situations are of major influence on the choices a traveller makes, either long-term or short-term. They have an effect on their learning behaviour, their habits, etc. Most research has focused on the situations that occur often, albeit these situations can show daily variation or unreliability. No specific research has been dedicated to find out which situations do not occur regularly or are unexpected, but can have a substantial impact from the viewpoint of a traveller or a road operator. Such situations can be called non-recurrent. Often traffic situations are viewed upon from a general point of view. As such, ordinary and daily traffic jams, or normal traffic flow without congestion is not to be labelled non-recurrent. This view point can for example be found in Emmerink et al. (1995). However, this does not include the individuals which experience such situations. Someone unfamiliar with daily and regular congestion has a difference experience than a person who is familiar. A complete description of non-recurrent traffic situations which includes both the general overview and the individual point of view is missing.

In order to focus on the effects of and preferences for traffic information for non-recurrent traffic situations, it is necessary to determine the different types of non-recurrence. This leads to the following research questions, which includes both objective and subjective aspects:

1. Which traffic situations are non-recurrent?
   (a) Which traffic patterns make a situation non-recurrent?
   (b) Which characteristics of a traveller make a situation non-recurrent?

1.2.2 Traffic information

Traveller information has been existing for a long period of time. Traffic information, as a subset of traveller information, has been available for a shorter amount of time (particularly the last 20-30 years, in the Netherlands). There is not just one purpose of traffic information as it can improve the comfort (or discomfort in case of a driver knowing he will arrive late) of drivers, improve the distribution of traffic flows over a road network as drivers are able to make "better" decision, etc. The effect of having a road network with a more robust traffic flow is one of the main interests of the traffic managers for applying traffic information, whereas an individual is just assumed to be interested in the information allowing him to make the optimal choices for his current trip. However, the preference for ***traffic information*** of drivers are expected to depend on factors such as the traffic situation (including his or her personal context), personal characteristics and attributes of traffic information. Based on these varying elements, the preference for traffic information is the second research objective.

This leads to the following research questions:

2. What traffic information do drivers prefer in non-recurrent situations?
   (a) Does the preference for traffic information vary over non-recurrent traffic situations, and if so, how?
   (b) Does the preference for traffic information vary over travellers (drivers), and if so, how?

   (c) Which attributes of traffic information are of influence on the preference for traffic information?
   (d) What are the effects of traffic information?

1.2.3 Route choice behaviour

Both traffic information and traffic situations have an effect on choices a driver makes. These choices range from operational (which lane to take), to strategic (where will I work, shop, etc?) and everything in between. These choices naturally have different time scales. In the research we focus on route choices for private drivers. The objective is to describe and model ***route choice behaviour*** in terms of routes and individual choices, during non-recurrent traffic situations in which different traffic information is provided to drivers. More specifically the effects of the combination of non-recurrent traffic situations and traffic information on route choice behaviour is studied, which is a necessary step in being able to predict route choice behaviour under these circumstances. Choices to switch between different modes of travel are excluded in this research.

This corresponds with the following research questions:

3. What are the effects of traffic information on route choice during non-recurrent traffic situations?
   (a) How can route choice behaviour be measured?
   (b) How can route choice behaviour be modelled?
   (c) Which attributes are important in route choice behaviour?

1.2.4 Scope

This thesis focuses on the combination of non-recurrent traffic situations and traffic information, both related to each other and to route choice behaviour. We only consider private trips, excluding commercial drivers and transport of goods. This determines a starting point for the research, but still covers a large area of interest. In order to make the research effort feasible, the scope of the research is further restricted in terms of methodology and content.

Firstly, it is assumed that most effects of traffic information in non-recurrent traffic situation are to be found in urban areas, both in terms of overall travel performance as well as individual experiences. Therefore this research is mainly geared toward road networks which include a large percentage of urban roads. This does not mean highway or rural roads are not of interest. For traffic information to have most effect it is important to have alternative routes available. This is even more important when a traveller encounters or is informed of a non-recurrent traffic situation, as at that time and location there has to be a possibility to change his/her route. If no alternatives are available at departure or en-route, the effects of traffic information will mainly be in the comfort of travellers. The most alternatives are found in urban areas, especially bigger cities. These usually have one or more ring roads which allow drivers to select either side and these rings roads have numerous options to divert. These urban areas also have one or more connections to a highway and would allow for local traffic to use the highway. Therefore, this research focuses on urban areas.

Second, the introduction of non-recurrence in both route choice and preferences for traffic information assumes the information about these traffic situations is readily available. This is increasingly becoming reality, as new technologies such as the tracking of GSM phones or bluetooth devices creates the possibility to measure travel times, speeds and routes in all areas. An example of such traffic information is TrafficHD by TomTom (Cohn, 2009). This research however, does not directly pay attention to the availability of the actual traffic information. In the future, further developments of existing technologies and complete new technologies using ad-hoc sensor networks, possibly the vehicles themselves, allow for more information on the traffic situation. Besides, a personal navigation device is expected to be able to determine the experience of a driver in a certain road network, and learn about the driver's destination. These technologies would improve the detection of varying non-recurrent traffic situations. In other words, we assume it is possible to distinguish between recurrent and non-recurrent traffic situations on an individual level.

Third, the research focuses on the traffic engineering side. This means we focus on aspects such as travel time, delay, lengths, routes and choices in a travel context. These aspects are all related to the traffic situations, traffic information and route choice behaviour. However, for traffic information, the interface and actual content of the information is also of importance. The difference of describing and prescribing traffic information has been studied (van Berkum & van der Mede, 1993) and show significant difference from a users' perspective. Next to the formulation of text, the user interface is important (Dicke et al., 2004). There is a vast amount of information on user interface design and usability (see (Wikipedia, 2011a,b)). However, this research is not about the impact of different interfaces and designs on choice behaviour. The elements which have an interaction with users have been carefully chosen, based on available literature and feedback in preliminary tests. Using this approach means it is not possible to relate the impact of traffic information or route guidance to the chosen designs and formulations, as there are no variations in these.

Fourth and last, it is known that traffic information both has a short term and long term effect on behaviour of travellers (see for example (van Berkum & van der Mede, 1993; Chen & Jovanis, 2003; Abdel-Aty & Abdalla, 2004; Bogers et al., 2007; Bogers, 2009; Ben-Elia et al., 2010). Drivers daily using the same route and experiencing traffic information or route guidance are able to learn about the network performance and the quality or reliability of the traffic information and route guidance. In this research we are only investigating the short term effects. This means that effects for a single trip are of interest. This is also due to the restriction of non-recurrent traffic situations. The nature of these situations, regardless of the type of non-recurrence, causes each trip to be "unique", which makes it more important to investigate the effects of information for the trip in which these traffic situations occur. This also excludes the learning of the performance (in terms of quality, reliability, etc.) of the traffic information and/or route guidance. To be able to investigate the learning behaviour of travellers regarding the information received, it is necessary to vary various aspects of the information, such as quality, reliability, accuracy, etc. This is not part of the research, as only complete and current information is provided.

1.3 Relevance

The main contribution of this research is to show the impacts of non-recurrent traffic situations and traffic information on route choice behaviour. The impacts are given for the overall routes and their characteristics, as well as individual choices. To be able to show such impacts, a framework for non-recurrent traffic situations is developed and the preferences for traffic information in these situations are determined. This leads to the following products/results:

- A framework for defining non-recurrent traffic situations, which is based on both the network and individual point of view. Such a definition enables researchers to distinguish traffic situations and possible effects of measures or behaviour.
- Insights in the preferences of drivers for different types of traffic information, concerning matters such as content, quality and price.
- A simulation environment (the AIDA RCS) which can be used for multiple research projects involving route choice behaviour and possibly also departure time and destination choice. As this environment is self-build, it is easy to adapt to other research problems. The AIDA RCS has already been adapted and used in a study on the effects of advising for safe routes combined with monetary incentives (Bie et al., 2011).
- Insights in the effects of traffic information on individual trips and route choice behaviour during a specific non-recurrent traffic situation (accidents).

This contribution is further discussed below in terms of relevance for science on one hand and society and practical problems on the other hand.

1.3.1 Scientific relevance

The main scientific contributions are:

1. a framework to classify non-recurrent traffic situations;
2. the preferences of road users for various kinds of traffic information;
3. a research tool for acquiring data on route choice;
4. a new type of route choice model, and;
5. insights in route choice behaviour during non-recurrent traffic situations with traffic information.

First, a new framework is created which includes demand, supply, and both endogenous and exogenous disturbances from the viewpoint of the road operator and the individual traveller. Such a framework was not available. The framework enables to distinguish various traffic situations and relate these situations to different behaviour and attitudes. In the framework, the traffic flows as a result of the transport system with its disturbances is used, together with the context of the traveller and his/her trip. The context provides extra information on the knowledge and previous experiences of a traveller which provides a better description of non-recurrence. The framework enables others to extent the investigation of effects that non-recurrent traffic situations have on the performance of the road network, as well as on the individual behaviour.

Second, non-recurrent traffic situations are shown to have a significant impact on the preference for traffic information. The framework provides a valuable way to define various traffic situations for which travellers prefer a different kind of traffic information. Such a specific attention to the effects of non-recurrence has not been found in relation to traffic information. The preferences also show the relation between the attributes of traffic information and its contents, for which the application of discrete choice models is shown to be an appropriate approach. As the preferences also differed between groups of drivers, for example based on their mobility type, it indicates that using traffic information is not easily explained. When investigating use and preferences for traffic information it is necessary to clearly define the groups and traffic situations of interest.

Third, the development of a simulation environment for route choice analyses (the AIDA RCS) allows for a large amount of studies to be undertaken. The AIDA RCS uses Paramics (a microscopic traffic simulation model) as the traffic model for vehicle movements, travel times, road network, etc. As such it differs from other route choice simulators which use mesoscopic traffic models, or use generated or real data on travel times or speeds. The microscopic approach enables detailed simulation of vehicle movements, as well as calculation of speeds and travel times which can be used for all kinds of traffic information or route guidance. The environment can be used locally, but also on the internet, which allows participants not to be physically present at the research facility. Apart from this, it enables studying multiple participants in a single experiment at the same time. This enables research to investigate effects as oversaturation, overreaction and concentration as presented by Ben-Akiva et al. (1991). Apart from route choice behaviour and various aspects of it (such as learning, day-to-day dynamics and navigation), other types of choice behaviour can be studied using the AIDA RCS. For example, the environment can be adapted to include departure time and destination choices. The flexibility of the environment makes it a useful addition to the toolkit of researchers interested in choice behaviour in traffic systems.

Fourth, a new type of route choice model is developed, which uses the exits of a junction as the alternatives a driver chooses from, independent of the choices at previous junctions. This type of choice model differs from other route choice models, as it reduces the choice set to the number of exits available, whereas for other route choice models it is necessary to generate and select relevant routes as the alternatives. This route generation and selection is a time consuming process, especially in case of en-route changes in a route. The proposed model, only needs to calculate a route to the destination starting at each exit. The exit choice model proved to be a suitable model for application in case of accidents and traffic information, and is able to predict the choices made by drivers with only small differences. Furthermore, the assumption of drivers making a choice at a junction independent from previous choices proved to be valid.

Fifth, using the proposed exit choice model we gained insights into the impact of traffic information on route choice behaviour in non-recurrent traffic situations. The impacts are shown on the overall route choice and traffic performance as well as individual trips and individual choices. Traffic information enables travellers to reduce their travel time regardless of the non-recurrent traffic situation they encounter, where especially guidance provides a large reduction. The traffic information causes the drivers to choose for the same routes, possibly leading to a concentration of traffic. Traffic information and route guidance proved to be very important aspects for choosing an exit, together with freeflow travel time, delay and the complexity of a route.

1.3.2 Societal and practical relevance

The findings in this research are relevant to public authorities and road operators. They apply traffic management in order to improve the overall performance of the road network. Specific goals are for example minimising the amount of lost vehicle hours on the highways, or achieving a certain amount of reliability for travel times. Safety can also be one of the main targets. In order to reach these goals, several strategies are applied, amongst which the provision of information to travellers. Traffic information used as such, aims for a better distribution of traffic on the road network, as drivers have a better knowledge of the locations of (and delay caused by) traffic jams. A better informed driver is assumed to make "better" decisions regarding his or her departure time and route choices. "Better" in this case refers to better in terms of network performance, which does not have to coincide with a better individual performance (such as smaller travel time). Information on accidents are usually communicated using various media. As this research shows, the provision of traffic information should incorporate the various traffic situations that drivers' experience. This includes the context of their trips, as being unaware of common traffic situations can greatly influence the preference for traffic information. Apart from the trip and its' context, the specific type of driver is of importance. Personalizing traffic information would enable road operators and road authorities to have a larger effect of the provided traffic information, in terms of their specific goals.

Next to the personalization of traffic information, this research also indicates the need to improve the detection of congestion and specific causes of congestion. Even though the technological advances for tracking vehicles provide a more complete picture of the current traffic situation, these technologies have difficulties tracking slow moving traffic in cities. In cities, traffic is often mixed with pedestrians, cyclists and other travellers. It is difficult to distinguish these different travel modes using just bluetooth or GSM tracking, even though matching the information with maps certainly improves this (for example, if the accuracy of the location is high, it is possible to distinguish road traffic from sidewalks). If the signal of a (smart)phone can be linked to a car, it is possible to improve the detection of traffic situations greatly. Especially the detection of incidents might help drivers in urban areas to improve their own experience and thereby reduce the decrease in network performance caused by an accident. Apart from the provision of traffic information in urban areas, a better detection of events also allows for many other improvements. This might be interesting for service providers as a means of extra sources for traffic information, when it is combined with specific devices used by drivers.

Personalization is also of relevance to service providers and manufacturers of personal navigation devices or navigation software. The variation of preferences for different types of traffic information, and different attributes caused by different situations and different types of drivers shows that for each trip it is important to know more about the driver and context of the trip to be made. Personalized information or route guidance improves the personal experience of the driver, and might cause a higher brand loyalty. If information is personalized and delivered in a timely manner, with a high reliability, it also indicates a rather high willingness to pay for this information. However, this depends on the type of driver. This means both service providers and manufacturers need to invest in diversifying their services and devices, to be able to target all groups and preferences.

1.4 Approach and outline

In this research four main methodologies have been used in order to provide answers to the research questions and reach the research objective. The first concerns literature reviews. For all three elements in the research objective, literature has been used to gather a sound basis for this specific research.

The second methodology is the use of a survey. This survey targeted the needs and preferences of drivers for traffic information, related to non-recurrent traffic situations. As an internet survey dealt with hypothetical situations, this survey can be labelled a stated preference survey, which combined the hypothetical situations with various alternatives with a set of attributes and attribute levels.

For the route choice research, a route choice simulator was developed as the third method. As the route choice simulator concerns simulation environments, the choices can be regarded as stated preference, but it is not explicitly a choice with predetermined options and attributes as is normally the case for stated preference studies.

The fourth and last used method concerns the application of discrete choice models to both the preferences for traffic information and route choice behaviour. In both cases it concerns no continuous variable for which a choice had to be made, but a select number of discrete options. In this case the discrete choice models have been restricted to those based on random utility maximization.

Figure 1.2 shows the outline of this thesis. Chapter 2 presents the framework on the non-recurrent traffic situations, using both the network and individual view. In addition, we combine these elements and propose a definition of non-recurrent traffic situations which is used in the following chapters.

Chapter 3 presents an overview of literature on traffic information in general, and more specific regarding the preferences for traffic information and the effects traffic

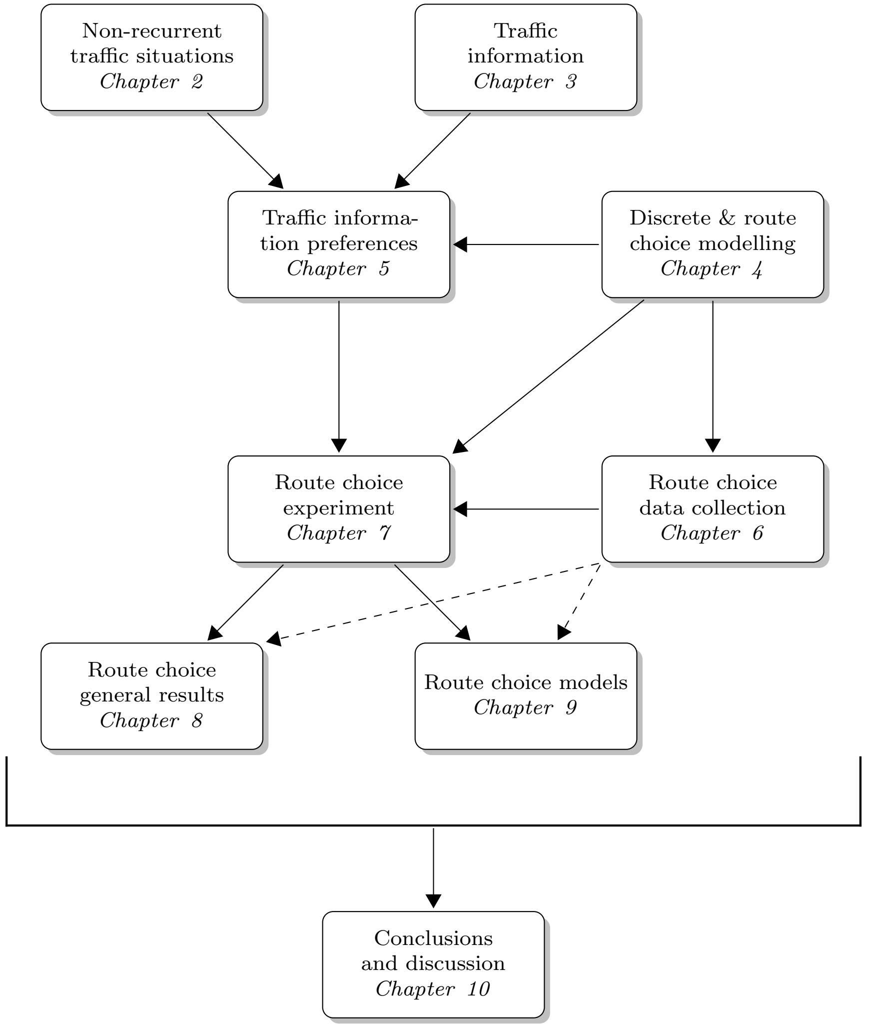


Figure 1.2: *Schematic outline of this thesis*

information has on route choice, both with elements such as content of information and the influence of personal characteristics. Chapter 4 presents the theory on discrete choice modelling based on random utility maximization. The specific application of these models on route choice is also discussed, including methods to acquire the necessary data for route choice modelling.

The results from chapters 3 and 4 are used in chapter 5. It presents the Internet survey that was undertaken to investigate the preferences of drivers for traffic information. This chapter presents the approach used in this survey, the results of the statistical analysis hereof and the choice models which have been estimated using the survey, regarding the attributes of traffic information.

Chapter 6 - 9 discuss the development, application and results of the AIDA RCS. Chapter 6 starts with a discussion of various methods that can be used for the acquisition of route choice data and describes the development and elements of the route choice simulator. Chapter 7 presents the setup of the experiment that has been undertaken, in which the AIDA RCS is used. The experiment is described in terms of the used road network, scenarios, participants and used procedure. Chapter 8 and 9 present the results of the experiment. The first part concerns the overall results, examining the effects of the variations in the scenarios on the trips and questions asked to the participants after each trip. Chapter 9 continues the presentation of the results, but focuses on the individual choices during the experiments. This chapter also presents the various route choice models which were estimated. Finally, the conclusions and a discussion of all results in regard to the research objective is presented in chapter 10.

Even if you walk exactly the same route each time - as with a sonnet - the events along the route cannot be imagined to be the same from day to day, as the poet's health, sight, his anticipations, moods, fears, thoughts cannot be the same.

—A.R. Ammons

Chapter 2

Classification of traffic situations

2.1 Introduction

Traffic situations are often depicted from the viewpoint of the road network or its operator. Congestion then can be recurrent or non-recurrent. Recurrence is defined as being repetitive; happening again and again. Even though congestion shows daily variation, the repetitive nature of a lot of traffic jams means that congestion can often be regarded as recurrent. However, several different types of events create variations (which can be repetitive) which can significantly influence the occurrence of congestion. Examples of such events are accidents or bad weather. The resulting congestion can be called non-recurrent, as these events are of a varying nature. However, using the causes for congestion limits the definition of non-recurrence, as the same traffic situation might be rated differently by different travellers. For instance, a commuter with a long history for a trip has a different perception of the road network and its traffic situation than a traveller who makes the same trip without the same experience. The experience of the individual is expected to play a significant role in rating a traffic situation to be recurrent or non-recurrent and the ways an individual deals with such situations. However, a viewpoint which includes both the viewpoint of the road network and the traveller is missing.

This chapter presents the resulting framework to classify traffic situations as recurrent or non-recurrent based on the viewpoint of individuals and the road network. The first research question *Which traffic situations are non-recurrent?* can be answered using this framework. Section 2.2 starts with a overview of congestion on a road network. Several causes are discussed as well as the effects of congestion, particularly in relation to travel time. Next, the individual traveller making a trip in a (congested) network is discussed in section 2.3. Especially unreliability and uncertainty are important elements which relate the current traffic patterns to the individual traveller, using his perception, experiences and familiarity. Combining both points of view leads to the framework which allows us to classify traffic situations with respect to their recurrence. This framework is presented in section 2.4, together with three examples. The chapter finishes with a summary in section 2.5.

2.2 Traffic and congestion

Congestion is defined in a dictionary as an abnormal or excessive accumulation, or over crowding (of traffic). In case of traffic flows, congestion can be defined as the phenomenon that arises when the input volume exceeds the output capacity of a facility (Stopher, 2004) and means an excessive use of that facility. Even though congestion is a world wide issue, it generally occurs on individual road segments and usually is not ubiquitous. A typical distribution of travel demand over a day shows two peaks; that of the morning and evening commute (see figure 2.1).

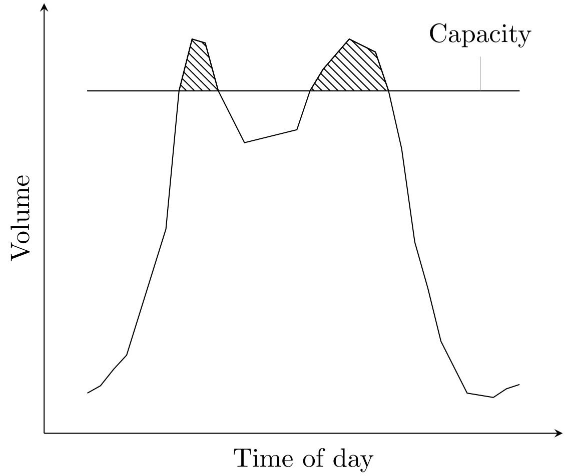


Figure 2.1: *Typical daily distribution of demand for travel (Stopher, 2004)*

This increase of demand is a common cause of congestion. The second cause is because of a bottleneck, which can be defined as a location where the capacity (supply) of a road segment is reduced (for example by reducing the number of lanes). Frequently, congestion is subdivided by its' cause into two categories: recurring and non-recurring. Recurring congestion is defined as the congestion caused by the routine or regular traffic volumes operating at regular capacity, or in other words, the congestion present on a normal day if nothing bad has happened on the road network (Hallenbeck et al., 2003). This would result in a "stable" traffic flow, as depicted in figure 2.2, excluding the shocks. Such a traffic flow is relatively stable, as there are always slight differences within and between days regarding demand (for example, differing composition of drivers and driver behaviour) which will cause different congestion. Non-recurring congestion is defined as unusual congestion caused by an event that is transient relative to the common or "stable" situation. These relate to the endogenous or exogenous shocks as given in figure 2.2.

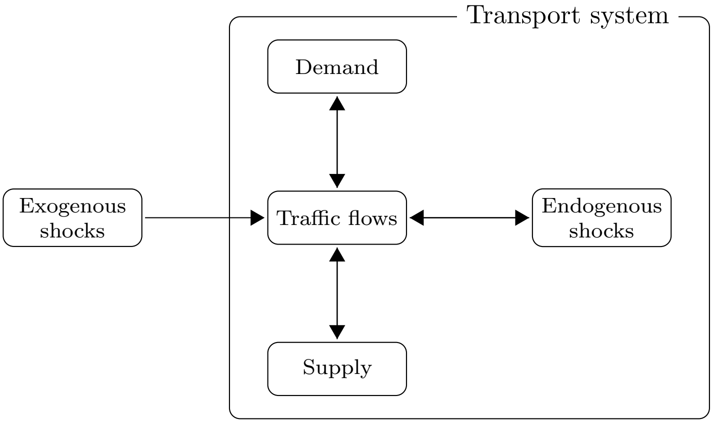


Figure 2.2: *Transport system with shocks (Emmerink et al., 1995)*

Several causes for congestion (regardless of being recurrent or non-recurrent) are often mentioned (Emmerink et al., 1995; Lomax et al., 2003; Hallenbeck et al., 2003) which are summarized as:

- incidents such as collisions, vehicle breakdowns and debris on the main road or just beside;
- road works;
- weather affecting the visibility or road surface;
- fluctuations in demand;
- special events, such as a concert;
- traffic control devices such as traffic lights;
- inadequate base capacity.

Another distinction is made by Wilmink et al. (2003). They distinguish between the regularity and predictability of a traffic situation. Each situation has its own characteristics. A resulting traffic situation can be caused by multiple sources, both predictable and unpredictable, and regular and irregular table 2.1.

The Dutch highway operator Rijkswaterstaat also registers the causes for congestion related to congestion weight (length of congestion multiplied by its duration). They use the following causes:

- temporary capacity reduction;
- high demand (including regular bottlenecks);
- incidents (such as demonstrations and people on the road);
- accidents;
- weather;
- road works;
- other.

The results for the past 11 years in relation to the congestion weight is given in figure 2.3. An excess of demand is clearly the most registered cause of congestion in 2010 (81%), followed by accidents (9%). Road works account for 7% of the total congestion weight and incidents for almost 2%. Non-recurring causes for congestion in total account for around 18% of all congestion weight.

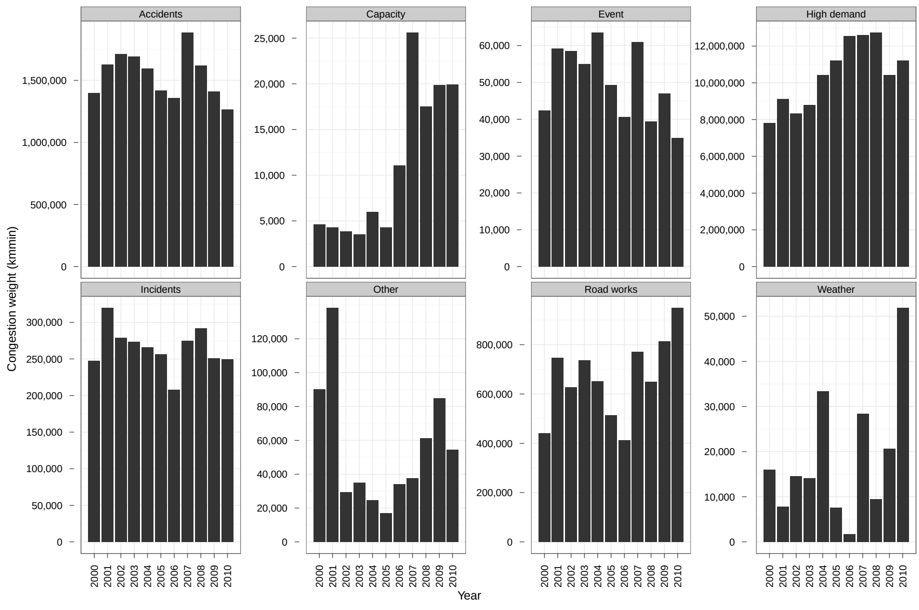


Figure 2.3: *Congestion weight in km min per year and cause*

Table 2.1: *Types of traffic situations (Wilmink et al., 2003)*

	Predictable	**Unpredictable**
Regular	Morning rush hour, evening rush hour, non rush hour, weekend traffic, bridge opening, small road works	Small incidents
Irregular	Holiday traffic, large events, large road works, special transport, weather	Large incidents, infrastructure failure, crises

Regardless of the cause of congestion, it has effects on travellers and their trips. Not all effects are negative. For example, Redmond & Mokhtarian (2001) show that commuters can perceive congestion as a welcome time to spend some private time in the car, enjoying music they like. However, the negative effects of congestion are more apparent. As congestion involves lower speeds and more variation in individual speeds, it has an impact on travel time, environment and costs (Stopher, 2004). Varying speeds directly cause an increase of fuel consumption and thus an increase of emissions and noise. This partly increases the costs for an individual to travel, but as the trip takes long, it also costs more time which might otherwise have been spent with a different activity (working for example). In this section we will focus on the effects of congestion on travel time.

Travel time and its key elements such as variation, reliability, predictability are often studied (see for example Noland & Polak (2002); Lomax et al. (2003); Tu (2008)). Travel time variability relates to the amount of variation or inconsistency in operating conditions, often caused by congestion. Such inconsistency can vary from day to day, within a day, from road segment to road segment and even from vehicle to vehicle. Variability does not mean an unpredictable travel time, but it relates to the uncertainty of a prediction for a travel time. Unpredictable variation of a travel time is caused by not knowing the laws which allows us to reproduce the actual situations. Recurrent congestion can still show variation, but this variation is predictable to a certain extent. Non-recurring congestion is often much harder to predict, but with information on the exogenous and endogenous shocks, it is possible to a certain degree. Other causes, such as accidents are much harder or impossible to predict (Hilbers et al., 2004).

Variability and predictability are both related to space and time (and aggregation level). Consider an accident on a highway, which means that the travel time for drivers upstream of this accident are likely to have an increase in travel time, which is unpredictable. Drivers entering the highway downstream of the accident have a possible decrease in travel time, which is also unpredictable. Within each group, the variability of travel time is expected to be low, but the travel time variability over the highway as a whole is increased. The same concept applies to road works, events, etc. Lomax et al. (2003) consider five levels of variation each with its own spatial and temporal level of aggregation:

- driver variation;
- road segment variation;
- time of day variation;
- daily variation;
- specific condition or event variation.

However, the previous discussion only concerns the highways. Urban traffic is expected to exhibit the same causes for congestion, but the amount of transient causes is expected to be higher, as there are many more events possible that cause a temporary or spatial variation in the demand or supply. An overview of urban traffic patterns is given in Weijermars (2007), which showed variations in patterns caused by seasonal and week of day effects. Also road works were an important factor in variation of urban traffic patterns.

2.3 Individual traveller

For an individual traveller it is important to have a reliable execution of a trip. Related to travel time, it means a traveller is interested in being able to predict the travel time for a trip with an specific amount of certainty. The amount of certainty is defined by the expected travel time, its' variation and stability and the information and alternatives available (Snelder et al., 2004). Reliability of door to door trips is considered a key objective for the Dutch Government (Ministerie van Verkeer en Waterstaat, 2005). Bates et al. (2001) mention two reasons for the importance of travel time reliability to travellers. The first is that travellers are sensitive to the consequences which they associate with variability such as late or early arrivals. The second explanation is that travellers value variability itself independent of its consequences, as variability increases anxiety, stress or additional cognitive burden with planning the trip.

The reliability of a trip is explained by Snelder et al. (2004) (see figure 2.4). The regular circumstances determine the normal, "stable" network performance, which

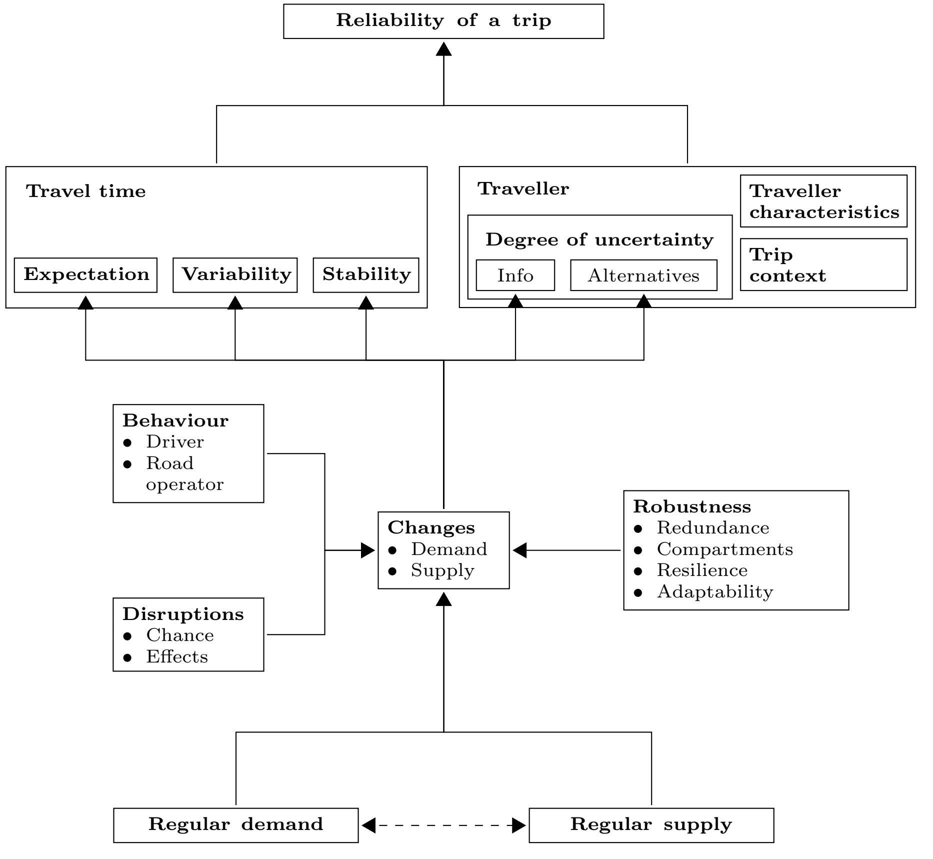


Figure 2.4: *Reliability of trips (Snelder et al., 2004)*

is affected by changes in demand and supply. This is a different relation as given in figure 2.2, where the changes are related to the resulting traffic flows, instead of demand and supply. These changes are determined by on one hand the robustness of the road network, and on the other hand by the behaviour and type of disruption. These changes in demand and supply have an effect on the travel time as such, and the degree of uncertainty which is experienced by a traveller. The effects on travel time can be described using expectations, variability and stability of the travel time. For the traveller, the personal characteristics and trip context are of influence on the degree of uncertainty, together with the information a traveller has acquired or received and the alternatives which are (known to be) available. For the trip at hand, these elements in turn determine the reliability.

Over the past years, reliability of travel time has been studied much (see for an overview Van Lint et al. (2008); Tu (2008)). Reliability is said to combine both predictability and variability of travel times, which can be quantified in various ways (Bates et al., 2001). Van Lint et al. (2008) provide a short overview:

- statistical range methods, which consider travel time windows using an expected travel time with a range based on the standard deviation;
- buffer times indices, which indicate the extra percentage travel time due to

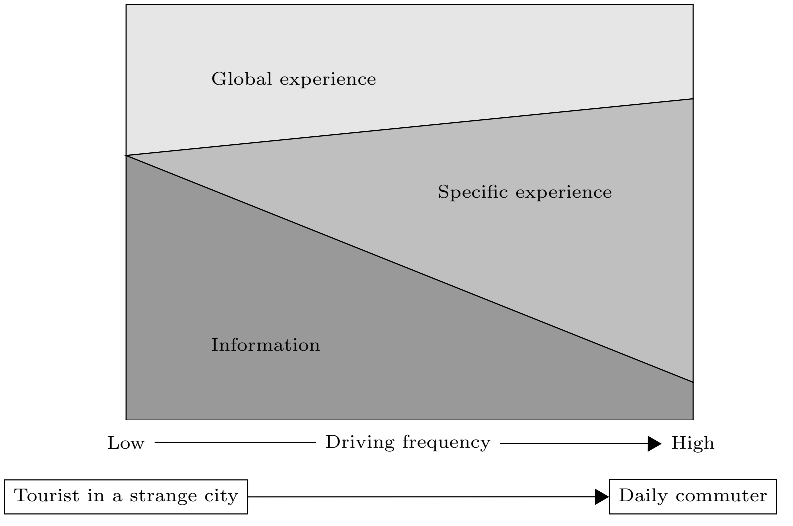


Figure 2.5: *Relative use of information and experience as a function of driving frequency (Stern, 2004)*

travel time variability on a trip that a traveller should take into account in order to arrive on time;
- tardy trip measures, which represent the travel time unreliability using the amount of trips that result in a late arrival;
- probabilistic measures, which express travel time reliability in terms of probabilistic measures, such as the probability that a trip can be made successful within a specific time frame.

However, they argue that neither of these measures is consistent and there is no definite conclusion regarding a measure for travel time reliability.

Regardless of the measure of reliability, they all relate the variability and predictability to expectations. Each traveller has its own valuation of reliability, depending on (amongst others) personal characteristics and trip purpose. As argued by Bonsall (2004b), a traveller bases a prediction on various sources of information:

- personal experience;
- second hand experience;
- information and advice.

Stern (2004) distinguishes two other forms of experience as provided in figure 2.5, relating these together with information to the driving frequency. Specific information refers to the individual's given surrounding road network, whereas global experiences refer to general structures of cities. The more frequent a traveller uses a road network, the more specific information one acquires.

Although Stern applies these concepts to just spatial information, they can easily be transferred to travel time. The more frequent one uses a road network, the more specific experience with the road network *and* its traffic patters a traveller acquires and uses. Such an approach is confirmed by, amongst others, Adler (2001);

Karl (2003). Driving frequency increases the familiarity of drivers and as such the ability to predict the variability (and reliability) of the travel times (and traffic patterns), which reduces the need and use of information. A traveller is able to be more certain when having more experience.

As it concerns perceptions, it is not possible to just use objective measures of reliability. Hilbers et al. (2004) distinguishes between the experienced reliability for those that have used a certain route or modality and the image reliability for those that have no real experiences with a specific alternative. Regardless of the modality, the image reliability is almost always lower than the experienced reliability. Especially those that do not use a modality perceive a low image reliability of a modality, with the largest difference for public transport. They argue that subjective reliability is mostly based on being in the hands of unforeseen circumstances.

In case of a trip in which a traveller faces uncertainty (due to his or her perceived unreliability), a traveller can use a number of strategies to reduce this uncertainty (Bonsall, 2004b):

- accessing additional information;
- advance planning;
- reducing consequences;
- accepting;
- capitalizing.

2.4 Framework

As already mentioned in Lomax et al. (2003), recurrent and non-recurrent traffic congestion does not fully describe the situations a traveller experiences. Putting focus on just congestion and its causes, diminishes the role of the traveller. In this section we present a framework which combines the concepts mentioned in the previous sections. This framework enables us to classify individual trips on a scale of non-recurrence, whereas the previously discusses frameworks have a different focus, or only classify causes of congestion.

The framework distinguishes the two main parts discussed in this section; the traffic system and the travellers. The traffic system can be described using objective elements, whereas the traveller is described using subjective elements. Both parts are related to each other using certainty and reliability, which are both objective and subjective. The relation of these parts and their elements is depicted in figure 2.6, which focuses on traffic. The physical lay-out of a road network is considered as input in the familiarity, perception and experience of the traveller, and also is input for the congested and uncongested network.

A road network exhibits a specific traffic flow as a result of the current demand and supply, with possible congestion. Such congestion can be described with the cause, which still can be labelled as non-recurrent or recurrent. The effects are also of importance, both of the congested and uncongested road networks. This effect determines the variability and predictability of travel time, besides the other

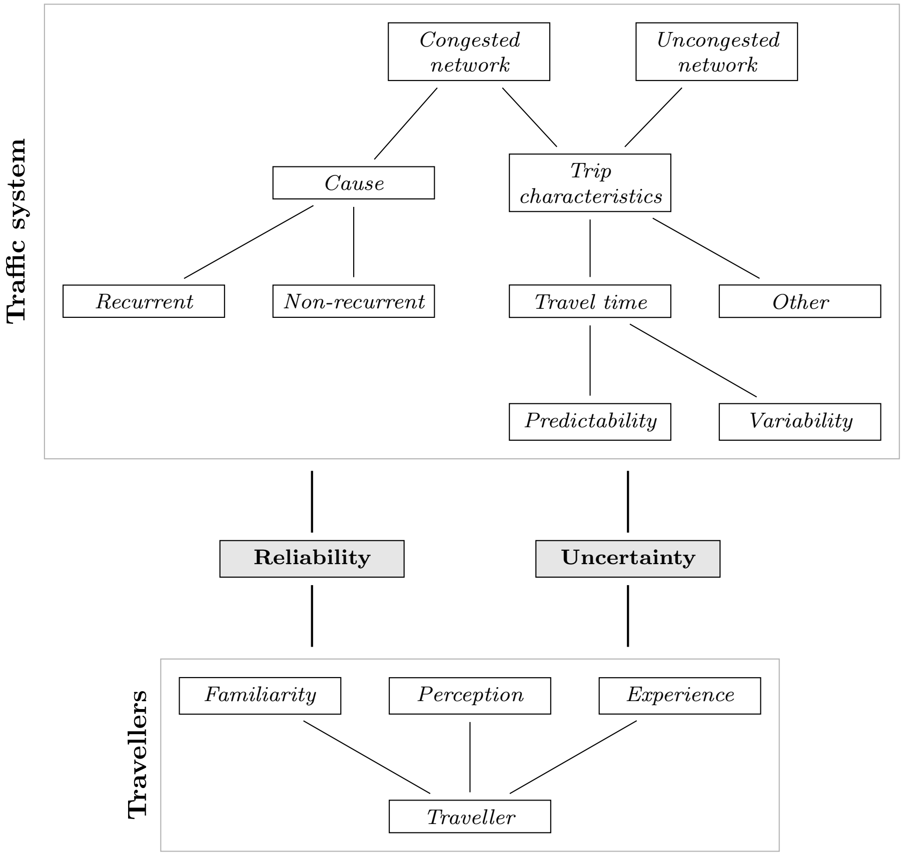


Figure 2.6: *Framework for non-recurrent traffic situations*

effects. The cause of the effect on travel time is also related to the predictability and variability, as an accident has different impact compared to road works or a bottleneck.

A traveller has a certain degree of familiarity and experience with a trip and the road network or transport system in which this trip is planned. This is combined with a perception of the road network (a mental map) and the traffic situations that can occur on this road network. The perceptions of the traffic situations range from very local knowledge (for example which lane is the fastest near a bottleneck) to an overall perception of the congestion at specific times, and the consequences for travel time and travel time variability (for example the notion that in the morning rush hour, a city is hardly possible to reach without major delays). As such, a traveller is able to predict the reliability of a trip, and a level of uncertainty herein.

Reliability and uncertainty relate the subjective elements (as considered by the traveller) of the traffic situation to the objective elements (as described by using the (un)congested road network and its cause and trip characteristics). Both sides have their own dynamics, but meet each other in the middle by a certain degree of reliability and certainty.

Based on the framework, we argue that a traffic situation should be used as a basis for describing as recurrent or non-recurrent in stead of congestion. A traffic situation is the traffic pattern a traveller experiences in the current trip. A traffic situation can then be defined as non-recurrent based on the traffic pattern with different causes of congestion and the traveller and his familiarity, perception and experiences. The degree of non-recurrence for a traffic situation is a continuous scale, as it is not possible to determine discrete levels for, for example, familiarity or experience. To answer the research question on which traffic situations are non-recurrent, all elements in the framework are needed.

To describe the framework, three examples that are of interest for this research are described in the following paragraphs. These three are:

- accidents;
- road works;
- large events.

2.4.1 Accidents

A large number of different types of accidents happen. They can be classified to for example the cause, the number of vehicles involved, the results of the accident (property damage, personal injuries or death), the number of lanes blocked, the duration, the road type where the accident happened, etc. For non-recurrent traffic situations with an accident not all of these classifications are important. Most important are the duration and the effect on the local capacity.

These two factors describe the impact of an accident on the congestion that is caused by the accident. Depending on the location of the traveller with respect to the location of the accident, the travel time a traveller will experience can be increased on decreased. In all cases, the location and time an accident occur is unpredictable, even though most travellers expect accidents to occur. The variability of travel time caused by the accident is thus increased. As a traveller is not expected to have specific experience for this accident, the uncertainty and unreliability are increased, making the traffic situation non-recurrent to a pretty high degree. This applies to all travellers influenced by the accident.

2.4.2 Road works

Road works are different from accidents in the sense that in most cases they are predictable, as road works are planned in advance by the road operator. This allows travellers to acquire information before their trip, which means they have a perception of the traffic patterns which will occur as a consequence of the road works. However, the specific congestion which will occur is not known at the start of the road works, but will increase with time in case of a traveller who makes multiple trips during the road works. The degree of non-recurrence for a traffic situations with road works thus depends on the experience of a traveller with the ongoing road works as well as the information that is acquired about the road works (apart from the familiarity with the “regular“ traffic situations and the road network).

2.4.3 Large events

Large events mean a different demand, often with travellers for a single location. The resulting traffic flow will be different compared to regular traffic flows, because these travellers are not all as experienced and familiar with the road network and regular traffic situations on this road network. Next, even though the travellers going to the event are aware of this event and can be expected to have acquired information regarding their trip accordingly, other travellers might be unaware of this event. Predictability thus differs between these groups of travellers. The variability of travel times for both however is expected to increase, based on the change of traffic flows. However, if a large event occurs more often, both type of travellers build an experience and are able to better predict the traffic patterns and their travel times. The degree of non-recurrence for a traffic situation with a large event depends on the traveller under consideration and the specific experiences and familiarity.

2.5 Summary

Congestion is a common product of the daily demand for mobility in an environment with limited supply. Especially during the peak periods, the capacity of the road network is insufficient, causing congestion. Congestion has sever impacts on the mobility of travellers and accessibility of locations, as congestion increases the travel time. However, not all congestion is similar. Congestion has different causes, which are often classified as recurrent or non-recurrent. Recurrent congestion is caused by routine or regular traffic volumes operating at regular capacity. Recurrent congestion is relatively stable, whereas non-recurrent congestion is transient, caused by exogenous or endogenous shocks. Examples of non-recurrent congestion are incidents such as collisions, road works or large events. In 2010, around 18% of the total congestion weight (in kilometer minutes) on the Dutch highways is caused by non-recurrent congestion.

Regardless of the cause of congestion, it has effects on travellers, both positive and negative. Especially travel times are of interest in this research. A possible positive effect is a longer trip which means a longer private time for a traveller. However, the negative effects are more prominent, as congestion increases travel time and travel time variability and thus the unpredictability of the travel time for a specific trip. Both variability and predictability are related to time and space as variation can exist for one trip and driver, a road segment or between time of day or days as a while.

For an individual traveller, the reliability or certainty of a trip is important. The degree of reliability or certainty of a trip depends on the variability of the travel time and the expectations of the traveller, which means that each traveller will have a different perception of the trip to make. Causes for congestion and the resulting travel time variation are therefore not the only information needed to describe the traffic situation of a traveller. The experiences and familiarity of a traveller with the road network and the traffic patterns that occur on this network and the perceptions of the traveller hereof are important in describing the reliability and certainty of a trip. As a traveller acquires more experience with a traffic pattern, he will be more certain about the expected travel time and thus he is able to predict the travel time with increasing reliability.

Combining both the traffic patterns, including congestion and its cause and effects, and the traveller with his familiarity, perception and experience, a new definition of non-recurrent traffic situations is given. A non-recurrent traffic situation is defined based on the traffic pattern with different causes of congestion and the traveller and his familiarity, perception and experiences, on a continuous scale. A traffic situation is used in stead of just a traffic pattern or congestion, and is defined as the traffic pattern a traveller experiences in the current trip. Non-recurrence can then be defined using the causes and effects of congestion, as perceived by the traveller. This makes non-recurrence a continuous scale, as more experience or a different perception leads to a different degree of non-recurrence. Using the framework it is possible to determine the need for different kinds of traffic information, as is expected that in case of uncertain or unreliability in a trip, there is more need for traffic information or route guidance, which depends on both the traffic patterns and the familiarity, perception and experience of the traveller. The framework is described using three situations with different degrees of non-recurrence:

- accidents;
- road works;
- large events.

The framework situation is further used in this research in chapter 5, in defining the elements of interest for the preferences for traffic information of drivers. Especially the three previously mentioned situations are used in this research, as these are quite common and can have a significant impact on travel times that drivers experience.

Information is not knowledge.

— Albert Einstein

Chapter 3

Types, characteristics and impact of traffic information

3.1 Introduction

Traveller information or Advanced Traveller Information Systems (ATIS) in general serves a number of important roles. It can (Lyons, 2006):

- make the individual traveller aware of the travel options available to him for a particular journey;
- empower the individual traveller to make more fully informed travel choices;
- assist the individual traveller in being able to successfully undertake and complete the journey.

This chapter focuses on traveller information specific for car drivers, which means that we focus on route- and departure time choice in particular and related choices such as destination (for example when searching for a parking place).

In general, the following factors are influential in affecting the preferences for, need for and use of traffic information and the effects hereof (adapted from (Wolinetz et al., 2001)):

- uncertainty - if there is little to be uncertain of in a trip (such as no variability in traffic conditions or a short remaining trip length) there is little need for traffic information;
- information awareness, access and use - travellers who are aware of available ATIS services and have access to these services are more likely to use them;
- situational and contextual characteristics - travellers with different trip purposes, or different flexibility in arrival or departure times have different preferences for traffic information;
- personal characteristics - travellers differ with respect to socio economic characteristics such as age, gender, income, risk attitude and learning behaviour and show different needs for and use of traffic information;
- information characteristics - the (perceived) characteristics such as quality of information affect the use hereof.

In this research we are interested in the use of ATIS in relation to non-recurrent traffic situations as given in research question 2: *What traffic information do drivers prefer in non-recurrent situations?* This chapter presents an overview of the available literature on traffic information and provides the basic elements necessary to answer this research question, which will be provided in chapter 5. We start in section 3.2 with an overview of the contents and types of ATIS. In section 3.3, the characteristics of traffic information which influence the preference for and impact of information are presented. An overview of personal characteristics which influence the preferences as well as the impact of ATIS is given in section 3.4, followed by a presentation of the possible impacts of ATIS in relation to the individual and the network performance. The chapter finishes with a summary in section 3.5.

3.2 Contents and types of traffic information

The amount of possibilities for the contents or types of traffic information are numerous. This makes it difficult to pinpoint the exact preferences of drivers for content. Van Berkum & van der Mede (1993) distinguish between descriptive and prescriptive information. Descriptive information provides information on the current state or lay-out of the network, either qualitative or quantitative. Prescriptive information tells a driver what to do. These types can very well be mixed, in which case the descriptive information can be used as a motive for the prescriptions. Schofer & Khattak (1993) also mention a difference between static and dynamic information. Traffic information can also be classified according to the importance of the information for an individual (Lyons, 2001):

- very important - information that is essential in determining the choice of mode, route or destination and whether or not to make a trip;
- useful - information that does not affect the the travel choices above but which enables a driver to make more convenient, comfortable trip with possibly greater confidence and assurance;
- irrelevant - information that offers no value to the traveller in planning or executing a trip.

The importance of and need for traffic information is found to depend on specific characteristics of individuals, such as social class, gender, age and existing use of information. Providing information that suits everyone is therefore a difficult, perhaps even impossible task.

It is also possible to distinguish the information with respect to the time it is received or acquired: pre-trip, en-route and post-trip. Each can have its own contents and can relate to different choices. For example, pre-trip information can be used to choose a modality or a destination; en-route information can concern the route to take, or the decision to switch routes, post-trip information can be used to update the perception of a road network to make a better choice in a new trip. ATIS can also be classified according to the modality (Kenyon & Lyons, 2003). They distinguish unimodal and multimodal. With respect to types of traffic information, another difference is the origin of the data (Toledo & Beinhaker, 2006). Traffic information can be based on either historic data, real-time data or predicted data (Ben-Akiva et al., 1991). Traffic information based on predicted data is for example an expected travel time for the complete journey, using the travel time at each segment of the trip at the time you are expected to be there. When using real-time data, the travel time is based on the current travel time for each segment. Chorus et al. (2006b) furthermore distinguish between the way the information is acquired before a choice is made. They distinguish between active acquisition of information as a traveller actively searches for information and passive acquisition where a traveller uses the information available. The information available can either stem from the active search or a service who may provide the traveller with information unasked for. The first concerns information pull, while the latter concerns mostly information push. An example of information pull is the acquisition of multimodal information to find the alternative modes to get to a destination. Information push can be the information provided via a radio service, or a variable message sign.

A combination of type of traffic information and contents of traffic is defined in AGV & ITS (1996):

1. descriptive congestion information (length, location, cause of congestion, etc.);
2. enriched congestion information (delay time, travel time, time of arrival, etc.);
3. descriptive network information (indicating one or more route alternatives);
4. enriched network information (indicating one or more route alternatives and the expected travel time);
5. advising network information (advice for the best alternative route).

Results from this study in 1996 showed that 58% of the respondents who wanted traffic information in their car (which was 66%), preferred the network information. Of this group, 59% showed a preference for the enriched type, indicating they want information that still allows them to choose by themselves. However, if more reliable travel time predictions had been available, more respondents would have chosen the enriched network information.

Relating the type of information to a choice between the usual route and an alternative route, Dia & Panwai (2007) found that drivers are most influenced to change their route by prescriptive information with over 50% indicating they would definitely take their alternative route. In case of qualitative information on unexpected congestion on the usual route, just 32% indicated they would definitely take the alternate route.

Several studies are available that have taken a more detailed approach to the contents of traffic information (Streff & Wallace, 1993; Wallace & Streff, 1993; Penttinen et al., 2003; Mensonides, 2004; Ministerie van Verkeer en Waterstaat, 2004; Chorus et al., 2007a). The early results from Wallace & Streff (1993) are presented in table 3.1. In this study a survey was undertaken with 2,764 respondents, in which these respondents were asked to give ratings to information bits, in three different conditions: (a) commute to work, (b) non-work trip in a familiar area and (c) non-work trip in an unfamiliar area.

The top ten of the information bits chosen in these three conditions is depicted in table 3.1. The information bits are rated on a scale from one to five, where one is not at all important and five is extremely important. The resulting ratings of all three conditions showed that there is no consensus among all respondents. There was quite a large range in the ratings of individual information bits, but no information bit is in the very important range. The amount of information bits

Table 3.1: *Preferences of road users concerning the subject of information (Wallace & Streff, 1993)*

	Commuter survey	Familiar area (Non-work)	Unfamiliar area (Non-work)
1	Delay time on original route	Condition of alternate route	Condition of alternate route
2	Predictability of delay on original route	Congestion on alternate route	Availability of directions for alternate route
3	Level of congestion on original route	Predictability of delay on original route	Congestion on alternate route
4	Problems that will arise if late	Level of congestion on original route	Repair of alternate route
5	Type of problem	Delay time on original route	Ease of access to alternate route
6	Condition of alternate route	Repair of alternate route	Level of congestion on original route
7	Congestion on alternate route	Type of problem	Delay time on original route
8	Ease of access to alternate route	Ease of access to alternate route	Predictability of delay on original route
9	Distance to problem	Time of day of trip	Time of day of trip
10	Time on alternate route	Distance to problem	Type of problem

with an average rating above 3 was around 20 (total number of information bits was 33). This indicates that respondents think most information to be important. The top ten, however, show some consensus. Based on the complete listing, the commuter and unfamiliar area condition are the most different. These results show that it is important to provide for a large range of contents for traffic information. It also indicates that a different need may exist in situations where the traveller is unfamiliar with the situation.

In a internet survey from 2000, Nee et al. (2001) investigate the preferences of drivers for traffic information. Survey results suggest that when planning a trip, warnings of incidents, construction, and special events as well as possible alternative routes can be a useful supplement to typical traffic measures which describe current conditions. The results also indicated that users prefer a format that presents a quick impression of general current traffic which does not require much time to evaluate the information presented. The preferred option is to get an overview from a color-coded flow map to obtain the big picture of traffic conditions.

A more recent study undertaken in Finland (Penttinen et al., 2003) concerned the importance of several telematics services and the willingness to pay for these surveys. Traffic information was seen as a part of these telematics services. The services related to car drivers were ranked on their importance as follows:

1. real time weather information (pre-trip);
2. incident information for drivers (pre-trip);
3. incident information for drivers (on-trip);
4. ISA (Intelligent Speed Adaptation) forcing for own car;
5. ISA warning for own car;
6. in-vehicle monitor (current traffic signs);
7. the fastest route by car (pre-trip);

8. the closest available parking lot;
9. car navigation;
10. prediction of driving time (on-trip);
11. prediction of driving time (pre-trip).

The results indicate that the respondents do not see a great need for navigation or route guidance. The most important type of information is the weather information, followed by incident information pre- and on-trip, which differs from the normal situation. However, car drivers who thought navigation to be important were willing to pay the most for this, compared to other services. Also incident information (pre-trip and on-trip) are valued highly. The respondents also think that for car drivers, most of the public funding should go to incident information, whereas car navigation should receive the smallest amount of public funding. Another interesting conclusion can be drawn from this study. The fact that real time weather information was ranked most important indicates a situation where changes in weather conditions can greatly influence traffic (snow, ice, etc.). This would probably be different in a area like Italy or California, where weather conditions might not be as influential. The importance of services is thus related to the local context of these services.

In 2004 the Dutch government published a quantitative study on traffic information, which aimed at extending insight into usage and preferences for traffic information (Ministerie van Verkeer en Waterstaat, 2004). The study focused on two types of car traveller: (a) commuters and (b) recreational traffic, heading to large events. Commuters have a more strict time schedule because of meetings at work, while the recreational travellers are more free in their time schedule, although it depends on the sort of event they are heading for. Both groups were offered four options for which they could indicate if there was an added value to the currently available information and the willingness to pay for this extra information. The results are shown in table 3.2 where the total for each group is the percentage of respondents that see the option as added value (with the total number of respondents being 150). Recreational traffic sees all the options more as added value than commuting traffic does. However, a greater willingness to pay is found among commuting traffic. For both groups, the advice for the best route which takes account of congestion was most often chosen.

The same study also investigated the current usage of information for commuters and showed that many sources of information are available. The high percentage of respondents mentioning the possession of a car radio equipped with RDS-TMC is

Table 3.2: *Added value and willingness to pay of Commuters and Recreational traffic (Ministerie van Verkeer en Waterstaat, 2004)*

Information option	**Commuting traffic**		**Recreational traffic**	
	Willing to pay	Not willing to pay	Willing to pay	Not willing to pay
Information on expected travel time	12%	35%	7%	52%
Information on delay on chosen route	11%	32%	6%	72%
Information on the best route, accounting for congestion	16%	44%	18%	67%
Being able to select your own preferred information	17%	40%	13%	61%

striking. The study considers this percentage too high, due to misinterpretation by the respondents (they do not actually have RDS-TMC). Of the commuters, for who the most important sources of information are teletext and the tv, 40% acquires information before they travel. But also the internet and the radio are important sources. The most important information sought is about congestion (location and length). During the trip, 25% always uses information and 45% sometimes acquires information. The most used source for this information is the car radio (88%) while the mobile phone with an information number comes next with just 9%. The VMS (Variable Message Signs) are used in 5% of the cases. Again the type of information they are looking for is on congestion. Interestingly, on the way back from work, most respondents use the internet as their information source. The respondents also suggest some points for improvement. The top three in this list are:

1. the information is currently inadequate;
2. the information on short and daily congestion is not provided;
3. the information is too general and not specific for my trip.

Recreational traffic was also investigated in the same manner. Compared to commuters, almost the same number uses information before the trip (42%), but in this case the internet is more often used as a source. Teletext and radio come second and third respectively. The sought information also changes: 50% is looking for congestion information, while 32% wants to know information on the car routes to use. During the trip, a lot fewer respondents use information compared to commuters (only 38%). On the way back, just 5% acquires information, probably due to difficulties finding the information.

The study also investigated the content of the information both groups would want to have, but in a qualitative manner. Commuters prefer to have information on their route, and the possible alternatives (personalized information). They would also like to have more detailed information on congestion and travel times. Commuters also prefer their traffic information to be personalized. They indicate that they want the information to be more current and more complete. The big difference between both groups is that the commuters want information that takes away the uncertainty, while recreational traffic is looking for information that takes away their unfamiliarity.

Another Dutch study aimed specifically at information for congestion caused by incidents (Mensonides, 2004). The study consisted of two parts: one for a general traffic information preference, and a second part for the incident situations. The general part of the study considers three choices for information contents for respondents, in order of importance. This resulted in table 3.3. The majority of the respondents in the survey that was undertaken preferred to have information on accidents, although a large group also choose daily congestion as a first choice. It is also interesting to note that a large group choose information on road maintenance as a second choice.

In the case of an incident, the preferred contents of information were surveyed as well (Mensonides, 2004). This resulted in table 3.4. The results show that people prefer information about length and location of congestion in the case of

Table 3.3: *General preferences of road users concerning the subject of information (Mensonides, 2004)*

	First choice	Second choice	Third choice	Total
Accidents	79	47	14	140
Road maintenance	18	58	42	118
Daily congestion	46	20	26	92
Events	4	12	31	47
Slipperiness	12	13	28	53
Fog	0	9	18	27
Total	159	159	159	

Table 3.4: *Accident specific preferences of road users concerning the subject of information (Mensonides, 2004)*

	First choice	Second choice	Third choice	Total
Length of traffic jam	38	18	16	72
Expected travel time	29	24	12	65
Alternative routes	15	39	45	99
Delay times	28	23	14	65
Cause of the incident	17	31	47	95
Location of traffic jam	32	24	25	81
Total	159	159	159	

an accident. However, when looking at the total, the alternative route was most often chosen. Also the cause of the congestion (more detailed information on the accident) was chosen quite often, but not as a first choice. This implies that information on length and location of congestion is not enough in most cases.

Chorus et al. (2007a) investigated the needs for information with respect to different information types, distinguishing between car drivers and transit users. Based on statements such as "I strongly need travel information concerning travel- (arrival-) and departure times for trips towards frequently visited destinations", respondents in the websurvey chose their agreement with this statements on a 5 point scale (1 is very much disagree and 5 is very much agree). The results with respect to the most-used mode are given in table 3.5.

Table 3.5: *Need for specific types of information (Chorus et al., 2007a)*

Information type	Frequently visited destination	Destination never visited before
Early warning function in case of disturbances	3.85	4.08
Full trip assistance	3.54	4.03
Time related information	3.46	4.54
Personalized information	3.38	3.89
Location specific information	2.98	3.43
Multimodal information	2.92	3.63
Cost related information	2.78	3.85
Information on than time and cost related aspects	2.75	3.57

In case a destination is never visited before (which can be labelled as a non-recurrent traffic situation) travellers indicate a largerd need for each type of information compared to frequently visited destinations. Next, travellers are more interested in information that makes choosing easier (such as full trip assistance or early warning) than in information that just provides more search options (such as multimodal information). In case of a frequently visited destination, the early warning function is most often selected to be of importance, which also shows a need for information that relates to non-recurrent traffic situations. With respect to unfamiliar trips, the need for traffic information is expected to be largest. En-route information in unfamiliar trips should make the execution of the trip as easy as possible.

The differing preferences with respect to traffic information caused by familiarity is also mentioned by Lyons (2006), although he argues that there may be limited demand for disruption alerts either pre-trip or en-route. Such information should be pushed to the traveller, which indicates the need for the service provided to know about the common or selected routes and modes of the traveller.

As summarized by Weisser & Horowitz (2002), drivers want to know if there is a problem, what the problem is, if there is a delay and how long that delay is, what the prescribed or alternate routes are and how these alternatives compare to the currently selected alternative. The resulting information that is needed is however dependent on the trip context and personal characteristics of the driver.

3.3 Characteristics of traffic information

Traffic information can possess a large difference in attributes, such as reliability and costs. In 1993 it was already mentioned that it was necessary to include attributes of traffic information when evaluating ATIS (Schofer & Khattak, 1993). Attributes that were mentioned are: (a) credibility, (b) reliability, (c) accuracy, (d) timeliness and (e) relevance. These kinds of attributes will be discussed in this section.

First of all, a general list of requirements has been found in a Dutch study (Ministerie van Verkeer en Waterstaat, 2003). The traveller wants:

- accurate information;
- information he can use;
- information suited to his needs;
- to be taken seriously;
- to be able to choose;
- an integrated complete story, not a small part of it;
- simplicity (with optional extras);
- information based on his knowledge.

Converting these requirements to attributes that are important for traffic information to be accepted and used (and perhaps paid for) results in a large list ranging from quality to personalization. Such a list is reproduced by many other studies. For example, Lappin (2000) reports results of a qualitative US study, in which respondents were concerned with:

- accuracy;
- timeliness;
- reliability;
- cost (capital and operating);
- degree of decision guidance and personalization;
- convenience (ease of access and speed);
- safety (of operation).

Reliability was expected to be the most important feature of traffic information (Lappin & Bottom, 2001). However, it is also one of the most difficult ones, for a number of reasons. Acquiring a reliable measurement of the current traffic situations is a difficult task, especially in urban road networks. Other information, such as cause of congestion, or expected clearance time of an accident location, is often not even available. Also communication and computation are limiting factors. Usability of convenience was also mentioned by (Golledge, 2002; Chorus et al., 2006a), referring to the fact that travellers appear to prefer information that makes travelling easier.

At first sight, reliability coincides with accuracy and timeliness. However, these attributes do differ. Accuracy concerns the precision of the information that is provided. For example, information on a traffic jam can be given as: (a) congestion between two major cities (inaccurate) or (b) congestion between starting at a certain location on the highway (km 56) with a certain length (7,6 km). Timeliness concerns the delay between the measurement of traffic on the road (for example, using loop detectors) and the delivering of information at the end user. Reliability defines if the information provided is actually correct. For example if the information on a traffic jam at a certain location with a certain length is the same as the road user experiences. These concepts of course can be translated to other types of information, or other modes.

Related to the attributes of the information itself are the attributes of a device that has to be bought to use information. These are also important. Using a nationwide survey with three user groups (dispatchers, private vehicle drivers and commercial vehicle drivers) Ng et al. (1995) analysed the importance of six attributes using a Likert scale of 1 to 7: (a) cost of the system, (b) accuracy, (c) timeliness of the information received, (d) type of visual display, (e) system dimension and (f) audio capabilities. All three groups showed significant differences over the attributes, where accuracy was considered to be the most important, followed by cost and timeliness. Audio capabilities were valued least important by all groups. Of these groups, dispatchers were willing to pay the most for ATIS (mean of $2017), followed by commercial drivers (mean of $400) and private drivers (mean of $275).

Relating attributes of the information to the source of information (Hato et al., 1999) indicated that drivers value the information from different sources in different ways. For example the scope of the information (related to the area covered in the information) means that traffic information with a specific scope (such as travel times on a Variable Message Sign or VMS along a route) are only useful in case a traveller can use this information (the routes on the VMS coincides with the route the driver is using).

Table 3.6: *Evaluation of traffic information aspects (Katteler & Broeders, 2002)*

Quality aspect	(really) bad		neutral		(really) good	
	At start	2 months	At start	2 months	At start	2 months
Reliability	22%	16%	35%	32%	43%	52%
Accuracy	23%	22%	40%	40%	37%	38%
Timeliness	44%	38%	36%	38%	20%	24%

In an experimental study in the USA with an in-vehicle route guidance system, Schofer et al. (1997) studied various features such as: (a) destination accuracy, (b) route quality, (c) startup speed, (d) route replanning and (e) button controls. The most important feature was the destination accuracy, closely followed by route quality, overall reliability and route planning performance.

Another experimental study in the Netherlands (Katteler & Broeders, 2002) also addressed certain attributes of an in-vehicle information system and traffic information. Three aspects were investigated shortly after starting and two months later. The results in table 3.6 show that reliability is considered to be good by most respondents, while timeliness is the opposite. After two months, the reliability has been increasingly evaluated as good, whereas accuracy and timeliness show only small or no differences.

Mehndiratta et al. (1999) also studied four attributes of traffic information, using a stated preference approach. The respondents were participating in a test with dynamic in-vehicle navigation systems. The studied attributes and levels were:

- update frequency: twice a day, several times an hour, every five seconds;
- prescriptive or descriptive messages: minimum time route recommendation or information on current delays;
- geographic coverage: main expressways only, expressways and arterials, door-to-door; and
- price: free, \$5/month, \$10/month (Seattle), \$15 or \$25/month (Chicago and Boston).

The results indicated that update frequency and geographic coverage were important to travellers. A majority of the respondents preferred the descriptive information. Respondents also indicated to be willing to pay at least some amount for the information, depending on the other attributes.

The costs of traffic information are an important factor with respect to the willingness to pay for traffic information (Schofer & Khattak, 1993). This addresses the costs and benefits involved in a decision to acquire traffic information as argued by Chorus et al. (2006b). The willingness to pay is likely to differ over types and contents of services and is likely to change in time, as reliability and thus demand increases. The willingness to pay for traffic information in general however, is low (Polydoropoulou et al., 1997; Kim & Vandebona, 1999; Wolinetz et al., 2001; Khattak et al., 2003; Zhang & Levinson, 2008). The willingness to pay increases with higher quality, accuracy or personalization of the information. It also depends on the personal characteristics of the driver; especially younger males (who are expected to be more tech savvy) are more likely to pay for traffic information. Travellers prefer to pay per acquisition rather than a flat monthly fee.

3.4 Personal and contextual characteristics

Most of the previous research also considered personal and other characteristics that influence the driver preference for ATIS. These characteristics range from socio-economic data, such as age, gender, level of education and income, to more advanced concepts of attitude towards mobility. This section will provide a short overview of findings in the available literature.

In general, the most likely users of traffic information are male, highly educated, "connected", high income travellers with longer than average commutes who face high rates of congestion (Petrella & Lappin, 2004). In the study of Mannering et al. (1995) several socio-economic characteristics were questioned in a nationwide survey on traffic and travel information. These characteristics were used to build an ordered logit model for importance of traffic information in an in-vehicle traffic information system. The resulting model showed that several characteristics influenced the importance, such as the average level of congestion experienced in commutes, gender, annual household income and number of vehicles per household. The same type of model has also been applied to the importance of visual display type and preferred ahead-miles on information on alternate routes and road conditions. These models also provide other important characteristics such as average number of days driving on highways, average commute time, level of education, marital status, age and whether the respondent is a flexible worker (no strict working hours). This already shows the vast number of possible personal characteristics that are important.

Other more conceptual characteristics of travellers are of importance in their tendency to acquire and use traffic information. For example, flexibility in schedules has been shown to be a significant factor in acquiring traffic information, where a lack of flexibility means a bigger need for traffic information (Polydoropoulou & Tsirimpa, 2010). Habit is also a very important factor in the decision to acquire and use traffic information (see for instance (Lyons, 2001; Gärling & Axhausen, 2003; Parkany et al., 2004; Bogers et al., 2005)). Habitual behaviour often means travellers do not consider other alternatives and thus are less likely to desire information about alternatives. However, as argued by Chorus et al. (2006a), this does not mean a traveller will never use traffic information. Contrary, in trips with a different context (departure time, destination, purpose) the traveller is likely to acquire and use traffic information, which might also affect the habitual choices.

Latent characteristics such as information processing capabilities greatly influence the use and effect of traffic information (Hato et al., 1999). This coincides with the need for easy to use information, as the capabilities of different travellers to actually understand and use the traffic information varies. Other latent characteristics of travellers are risk attitude and perception of alternatives (Bonsall, 2004b). For example, as travellers have a better perception of a road network and alternatives, they change the preference for information from prescriptive to descriptive (Adler & McNally, 1994). A perception of high reliability of the often used alternative also means a traveller is unlikely to change his behaviour when receiving traffic information (Abdel-Aty et al., 1997). Travellers can further be described as being risk-averse or risk-seeking, but depending on the trip and the information received,

Table 3.7: *Accident specific preferences of road users concerning the subject of information (Barfield et al., 1991)*

Cluster	Description	%
Route Changers	Willing to change route while and before travelling	20.6%
Non Changers	Not willing to change at all	23.4%
Route and Time Changers	Willing to change both route and departure time	40.1%
Pre-trip Changers	Willing to change departure time, mode and route before travelling, but not willing to change while travelling	15.9%

a traveller might reverse his risk attitude (Katsikopoulos et al., 2000). Ben-Elia et al. (2010) indicate that less accurate information means that travellers prefer a more reliable route and move towards being risk-averse.

Another important factor in the continuous use of traffic information is learning behaviour of travellers (see for example (van Berkum & van der Mede, 1993; Chen & Jovanis, 2003; Avineri & Prashker, 2006; Bogers, 2009; Ben-Elia & Shiftan, 2010; Lu et al., 2011)). This learning concerns getting to know the dynamic nature of the transport network with the help of information services, as well as getting to know the potential of these information services to support choices in this network. However, as mentioned by (Golledge, 2002), traffic information, especially prescriptive information might hinder the learning behaviour of travellers, as the information on the road network is transferred from the traveller to the service provider. This might make the traveller dependent for decision making on the service provider.

To describe the user of traffic information, or traveller in general, Barfield et al. (1991) and Conquest et al. (1993) identify major groups of travellers. In a survey they asked respondents how often they are influenced by traffic information regarding their departure time, mode choice and route choice. These results were used in cluster analysis to identify major groups of commuters. This resulted in four groups (see table 3.7) which then were analysed with other socio-economic data of the respondents, such as age, gender, income, availability of information sources and current usage of traffic information. However, most of these personal characteristics were of no significant influence. It is thus not possible to use these characteristics to say to which group a traveller belongs. The willingness to change is therefore an extra description of the preferences for traffic information. The source of information, number of sources used and frequency of using information has confirms this finding that there are different groups with a different effect on decision making (Khattak et al., 2008).

Attitude in general has shown to be another descriptive of individuals (see for example (Parkany et al., 2004; Anable, 2005). One of the approaches is to use the attitude towards mobility as a descriptive for the preferences (and possible effects) of traffic information. Such a concept is equal to the concept used for willingness to change departure time, mode and route, but on a higher level. Such a concept has been developed by Need (2002) for the Dutch Ministry of Transport. The concept defines five types of travellers based on their attitude towards mobility and a group which does not belong to one of these attitude groups (11% of the population).

These five types are:

Accepters They see mobility as necessary evil that is part of life. Travelling is not pleasant for them, but they do not actively search for options to improve their travel. Choosing a mode of travel is based on ease and practicality. This group is 15% of the Dutch population.

Deliberates They see mobility as a functional method of getting from A to B. They do not enjoy travelling, because they see travelling as a waste of time. They thus try to limit and optimize the amount of travel as much as possible. Their obvious need for certainty influences their choice of modes. They have to be sure to arrive on a certain time. This group is 26% of the Dutch population.

Conscious They see mobility as more than a method of getting from A to B. Travelling is a goal in itself to them, for their physical and/or emotional needs. They choose a mode that fits best to their current needs, such as destination, company and weather. This group is 17% of the Dutch population.

Competitors They see mobility as a symbol of independence, the possibility to go whenever and wherever they like. It is an expression of their social status. Choosing faster routes or work during travelling shows that they are "better" than others. This group is 14% of the Dutch population.

Enjoyers They see mobility as the method to visit friends. They especially see the social function of mobility. They enjoy being on the road and deal with time in a relaxed manner. They are good at putting congestion in perspective. Their choice of mode is merely based on the absence of stress. This group is 17% of the Dutch population.

The concept of attitude towards mobility has been used as a variable in two studies (Ministerie van Verkeer en Waterstaat, 2004; van den Broeke et al., 2004). In the study by van den Broeke et al. (2004) the distribution of all types is significantly different from the study by Ministerie van Verkeer en Waterstaat (2004). More respondents were defined as Deliberates and Conscious. This is due to the setup of the survey. In the study by van den Broeke et al. (2004) the respondents could indicate which type they were, in contrast to the study by the Ministry of Transport. In the study, for each type, a hypothesis was developed on the need for traffic information. Conscious travellers probably acquire different types of information before travelling. Competitors are expected to have the most need for information concerning car travel. Accepters have a need for information on the fastest and most efficient way to travel. Deliberates are expected not to differ much from the Accepters, and need information on efficiency of travel options. Enjoyers, as last, are expected to need information on uncertainty in their trip, and information on alternatives in case of disturbances.

In the other study (Ministerie van Verkeer en Waterstaat, 2004) a qualitative and quantitative research has been done for the need for traffic information. Not only was the attitude towards mobility used in this study but also a difference between two important motives; commuting and recreational traffic. The main conclusions were that commuters use more traffic information and are more willing to change their routes. However, the actual preferences for traffic information between both motives show a large overlap, but the motive for using traffic information is more important than the attitude towards mobility.

The motive for travel was also used as explanatory variable in the study of Mensonides (2004). Commuting, business and other motives showed a different preference for contents of traffic information. Apart from motive, familiarity with the road network and frequency of road use were used also. Preferences for all characteristics were deemed to be important. Familiarity was also found to be important for route choices in a study by Lotan (1997). In this study, familiarity was defined as familiarity with the road network and familiarity with the information system. Familiarity with the information system, or the traffic information itself, is especially important, as also suggested by Karl & Bechervaise (2003). Apart from familiarity with traffic information and the road network, they also argue that familiarity with the traffic situations are important. The familiarity relates to the knowledge and availability of route alternatives, which is also found to be an important factor in acquisition and usage of traffic information (Adler & McNally, 1994; Ramming, 2001; Peirce & Lappin, 2004; Chorus et al., 2007a).

3.5 Effects of traffic information

Effects of traffic information can be separated for the individual and the road network. With respect to the individual we distinguish effects for the current trip (short term) and over multiple trips (long term).

Chorus et al. (2006b) present a framework for the acquisition of information and the update of perception with respect to long term effects. It is based on four mechanisms:

- increasing the awareness of alternatives;
- altering the perception of the characteristics of alternatives;
- influencing the perception of the trip context and complexity of the choice situation;
- influencing the perception of the degree of completeness of the traveller's knowledge.

In case of long term effects of traffic information, Khattak et al. (1993) already mentioned that ATIS may change driver preferences over time, making them more risk seeking. Abdel-Aty & Abdalla (2004) show that familiarity with the device that provides the traffic information induces a higher degree of compliance with both pre-trip and en-route advice. In case a traveller has no long-term experience with the travel times and their distributions (a lower familiarity), the effect of traffic information is more positive (Ben-Elia et al., 2007). The also conclude that initially, traffic information increases risk-seeking behaviour.

Bogers (2009) defines long-term effects to be the effects on a traveller's expectations and perceptions. Even if a traveller does not choose a route, he may use this information to adapt the overall expectation of the travel time and its' distribution on this route. En-route and post-trip information influenced the expected travel times. van Berkum & van der Mede (1993) found that without en-route information, travellers hardly updated their perceptions of travel time. Once a VMS was used to provide en-route traffic information, they updated their perceptions by giving more weight to recent experiences. A traveller is also able to update the knowledge of alternatives in the network, which is also mentioned by Chorus et al. (2006b). However, as Golledge (2002) mentions, ATIS and especially guidance might also influence the cognitive model of the road network and the ability of a traveller to choose a route without information.

In general the effects of ATIS on the current trip (short term) are that the traveller has made a better informed and more conscious decision and reduced travel time and stress (Tsirimpa et al., 2007). The travel time under provision of traffic information is reduced, depending on the type of traffic information. Abdalla & Abdel-Aty (2006) find a reduction of up to 44% of the overall travel time when drivers are equipped with both pre-trip and en-route advice. In an overview of the possible travel time savings by Toledo & Beinhaker (2006), it is clear that there is little agreement on the possible reduction in travel times, as this ranges from 1% to 40%, depending on the type of traffic situation (recurring congestion, incidents, average trip length, size of road network, etc.) and type of traffic information. Depending on en-route and post-trip information, Bogers (2009) finds a reduction of maximally 10% over multiple trips. In case of traffic information which reduces the uncertainty caused by travel time variation, Ettema & Timmermans (2006) find that day specific travel time information causes a reduction of scheduling costs of a maximum of €1.

With respect to non-recurrent traffic situations, Tsirimpa et al. (2010) find that travellers are more likely to change routes if information on an incident or road closure for the habitual route is provided. As mentioned by Ben-Elia et al. (2007); Ben-Elia & Shiftan (2010), non-recurrent congestion are usually situations with a high degree of variability, which makes it difficult to provide detailed and accurate travel time information. Risk-seekers are more likely to respond to such information by seeking new routes. Levinson (2003) notes that the largest travel time savins are possible in situations with non-recurring congestion. However, the actual behaviour of travellers during non-recurrent congestion is not detailed.

The short term (during a single trip) effects of traffic information are often studied based on the compliance of drivers or their switching behaviour (see for example (Mahmassani & Liu, 1999; Chen & Jovanis, 2003; Srinivasan & Mahmassani, 2003; Chorus et al., 2009)). Compliance with advice is generally very high, for example 89% in the study of Chen & Jovanis (2003) using a simulator. In case of an advice for straight ahead, the compliance even is 94%. Based on a survey of real behaviour, Tsirimpa et al. (2007) find that 36% of travellers made changes during their trip, of which most were performed to avoid congested areas in their regular route. Using another travel simulator, Ben-Elia et al. (2010) find a compliance of 82% for high accuracy information, and 59% for low accuracy information.

The transport system as a whole can describe the effects of ATIS. As presented in the overview of Toledo & Beinhaker (2006), travel time savings are a very important effect, but traffic information also reduces the travel time variability. The effects do not just apply to individuals who use traffic information, but also show an effect for those who do not. For example, Emmerink et al. (1995) show that the travel time of uninformed drivers reduces, although not as much as that of the informed drivers. The overall reduction does however depend on the market penetration of the traffic information and type of traffic information. In case of post-trip information (information which is provided to a traveller after arriving, for example on the travel times of the chosen route and one or more alternatives), informed drivers only benefit if the market penetration is below 20%, whereas for en-route information, for most levels of market penetration the informed driver benefits, with up to 15% for the informed driver and a reduction of average network wide travel time up to 5% depending on level of congestion (Emmerink et al., 1995). Increasing levels of market penetration reduce the extra travel time savings. Enrique Fernández L. et al. (2009) show that a 50% market penetration of ATIS, already 90% of the gains have been achieved. This is also found in much earlier work by Mahmassani & Jayakrishnan (1991) where they find little additional system wide benefits with market penetrations between 25% and 50%.

However, the positive effects of ATIS have been questioned (Hall, 1993; Ben-Akiva et al., 1991; Arnott et al., 1991). This is very easily explained as when drivers are all guided to the fastest routes, these fastest routes are likely to be overloaded. Ben-Akiva et al. (1991) distinguish three possible negative effects:

- oversaturation - when a driver is unable to process the (large) amount of information in order to select the optimal alternative;
- overreaction - too many users respond to the information, causing some roads to be busy and others to be quiet, continuously changing;
- concentration - without information, perceptions of alternatives and their characteristics differ between travellers, leading to different travel choices. Information increases uniformity of these perceptions and lead to similar choices and possible higher levels of congestion.

Although the results of overreaction and concentration might be similar, the nature of these effects is entirely different. Overreaction is caused by drivers or information providers that are unable to accurately predict the choices of all travellers, whereas concentration is an effect of having better informed travellers. Overreaction can be prevented by having a more accurate and high quality information, whereas concentration can only be avoided if the traffic information specifically directs travellers over various routes in case of an equal origin and destination.

The underlying assumption in these potential disbenefits is questioned by Hall (1996). He argues that traffic information is not to be used as a method to achieve system optimal traffic distributions, but should be treated as a service to travellers to first improve their confidence and comfort and second to try to improve their perception of alternatives and characteristics which might lead to a better use of available network capacity.

3.6 Summary

Traffic information is a widely studied subject with a vast amount of publications. In this chapter we presented an overview of the literature on traffic information where we distinguish between:

- the content and type of traffic information;
- the characteristics of traffic information;
- the personal and contextual characteristics related to using and acquiring traffic information;
- the effects of traffic information.

These items relate to research question 2 on the preference of drivers for traffic information during non-recurrent traffic situations and its subquestions.

The overview of different types and contents of traffic information provided varying classifications for traffic information, related to time, source and content. Drivers show the largest influence of traffic information when this information provides them with an advice, which is prescriptive information. However, if predictive information was available and reliable, travellers would prefer such kind of information. The content of the traffic information which is preferred greatly differs, although information about disturbances are highly preferred by most. In general information related to travel times is important to travellers. From section 3.3 it is clear that also various attributes are important for traffic information. Examples are costs, reliability and accuracy. However, in case of non-recurrent congestion, the reliability and accuracy of traffic information might be perceived to be low by travellers. When information is found to be more accurate or personalized, travellers indicate they are willing to pay more for acquiring information.

A large range of personal and contextual attributes are of importance in the acquisition and use of traffic information. In general, a well educated, high income male with a longer commute is most likely to use and acquire traffic information. However, depending on the context of a trip (for example, trip purpose), others might be more inclined to use traffic information as well. Also familiarity is found to be important in the acquisition and use of traffic information. Several latent characteristics of travellers, such as risk attitude, habits and attitude towards mobility are also of influence on the use of and preference for information.

The effects of ATIS are discussed for the individual and the road network. Individuals are likely to show both short term and long term effects of using traffic information. The short term effects indicate a willingness to change routes (and even modes or destinations) and reduce travel times by using traffic information. In the long term, traffic information enables travellers to update their perceptions of the network and its characteristics. On the level of the road network, also travel times savings are found, ranging between 10% and 40%. Also travel time variability is reduced. The largest effects of traffic information are expected in situations where travellers are unfamiliar (or have no long term experiences) or where non-recurring congestion occurs. The positive effects of traffic information might be questioned, as it is possible that traffic information causes extra congestion as travellers are concentrated on the same routes.

However, no literature was found which specifically addressed several situational characteristics together with preferences for contents of traffic information. Most research was dedicated to one-on-one relationships between preferences and situations. The combination of personal and situational characteristics for non-recurrent traffic situations when looking at preferences for traffic information are noticeable by their absence. In case of effects on non-recurrent congestion, it was found that more insights are necessary that show the reactions of travellers when receiving and using traffic information.

The findings in this literature review are used in the research for the preference for traffic information (chapter 5), as well as the impacts of traffic information on route choice behaviour (chapters 8 and 9). The focus in these chapters is on the non-recurrent traffic situations, as presented in chapter 2.

Nothing is more difficult, and therefore more precious, than to be able to decide.

— Napoleon Bonaparte

Chapter 4

Discrete choice modelling

4.1 Introduction

Decision making, or choosing between alternatives, can be approached from three different angles (Van Berkum & van der Mede, 1993; Suhonen, 2007):

- normative;
- prescriptive;
- descriptive.

Normative approaches express how a decision maker should behave when facing a decision. The prescriptive approach is about helping the decision maker in making good or better decisions. The descriptive approach is about how the decision maker chooses (either rational or not rational) in real life. For this research we are interested in an approach that describes the way an individual chooses. This applies to both the preferences for traffic information as the choice for a route. In the case of choosing for different types of traffic information and in the case of a route choice it both concerns discrete alternatives, contrary to a continuous alternative such as petrol. One of the theories applied to discrete choice modelling in transport related cases is random utility maximization. This theory is chosen as it allows to investigate the importance and significance of various alternatives and attributes which can be easily related to real life situations. The relations and behavioural implications are more important in this research. Next, random utility maximization is a well described field in transport modelling and software to estimate models is readily available.

This chapter starts with the basic theory behind random utility maximization as it is applied to discrete choice modelling (section 4.2). This section starts with the multinomial logit model as the basic model of this choice theory. This model however has properties which are often violated in real life observations (section 4.2.2). To this end, recent developments which relax the assumptions leading to these properties, are described in section 4.2.3 to section 4.2.5. Route choice modelling is a particular problem in discrete choice modelling, and is described in section 4.3. This section describes the necessities for a route choice model: 1. path observations, 2. choice sets, and 3. choice model. Next, the proposed route choice model is given in section 4.3.4, followed by a brief overview of the important factors in route choice. The chapter finishes with a summary (section 4.4).

4.2 Discrete choice modelling

Discrete choice modelling in the context of traffic and transport has been much studied and described. We refer the reader to Ben-Akiva & Lerman (1985); Louviere et al. (2000); Train (2009) for an in depth presentation. Discrete choice modelling is based on a set of general assumptions regarding the following aspects:

- the decision maker and its characteristics;
- the alternatives and their availability;
- attributes of the alternatives defining the (dis)benefits of an alternative;
- the decision rule as the process by which the decision maker chooses.

An individual confronts a set of discrete alternatives, also known as the choice set, from which he or she has to choose. A discrete choice set contains a finite number of alternatives that can be explicitly listed. The identification of the list of alternatives is a complex process usually referred to as choice set generation. For the decision rule, it is assumed that the decision maker maximizes his or her personal benefit or minimizes the disbenefit. The (dis)benefit is described in terms of utility, where each alternative is given a utility based on (measurable) characteristics. However, a researcher cannot observe or measure all characteristics or experiences measurement error, which all mean a part of the utility of an alternative cannot be described. In random utility theory, this is translated in a utility function (equation (4.1)).

$$U_{in} = V_{in} + \varepsilon_{in} \tag{4.1}$$

in which:

$$\begin{aligned} U_{in} &= \text{Utility for alternative i and individual n} \\ V_{in} &= \text{Systematic utility for alternative i and individual n} \\ \varepsilon_{in} &= \text{Random or unobserved utility for alternative i and individual n} \end{aligned}$$

The systematic or observed utility is defined as a function of attributes of the alternative i and decision maker n and taste coefficients $V_{in} = f(\beta, x_{in})$. The probability of choosing alternative i from choice set C_n by decision maker n is given by:

$$\begin{aligned} P_n(i) =& Pr\,(U_{in} > U_{jn})\,\forall j \in C_n, i \neq j \\ =& Pr\,(U_{in} - U_{jn} > 0)\,\forall j \in C_n, i \neq j \\ =& Pr[V_{in} + \varepsilon_{in} > \max_{j \in C_n, j \neq i} (V_{jn} + \varepsilon_{jn})] \end{aligned} \tag{4.2}$$

It is often reasonable to specify the observed part of the utility function to be linear in parameters, together with a constant: $V_{ni} = \beta x_{in} + k_i$ in which k_i is a constants that is specific for alternative i. This alternative specific constant (ASC) denotes the average effect of the unobserved attributes on the utility. This causes the unobserved part of the utility function ε_{in} to have a mean of zero. As only the differences in utility matter, only the differences in the alternative specific constants are of interest (Train, 2009). To this end, the researcher must normalize the constants, by arbitrarily setting one of them to zero, which makes the remaining constants relative to the normalized alternative. The same kind of issue is at hand for socio-demographic variables, as these attributes (of the decision maker) do not vary over the alternatives. Again, for one of the alternatives, the attributes for the decision maker are normalized to zero, which means that the estimated coefficients for the remaining alternatives denote the relative effect of a socio- demographic variable to the normalized alternative.

The values of the attributes can be defined in two ways. The first way uses the actual value of an quantitative attribute, such as costs or travel time. The second way uses a selected type of coding, for qualitative or quantitative attributes. Possible attributes are comfort, reliability, etc. There are several ways of coding these attributes, of which more details can be found in Louviere et al. (2000) and Hensher et al. (2005).

4.2.1 Multinomial Logit Model

Based on the given approach it is possible to estimate the relative importance of various attributes of various alternatives. The estimates are provided by the $\boldsymbol{\beta}$. By assuming a certain distribution for ε_{in} it is possible to estimate these coefficients. The distribution most used for this purpose is the Gumbel or Extreme Value Type I (EV) distribution, which leads to the Multinomial Logit Model (MNL) proposed by (McFadden, 1974). If ε_{iq} is Gumbel or EV(η, μ) distributed with $\mu > 0$, then the probability density and cumulative density functions are given by:

$$f(x) = \frac{1}{\mu} e^{-\frac{x-\eta}{\mu}} e^{-e^{-\frac{x-\eta}{\mu}}} \tag{4.3}$$

$$F(c) = \int_{-\infty}^{c} f(x)dx = e^{-e^{-\frac{c-\eta}{\mu}}} \tag{4.4}$$

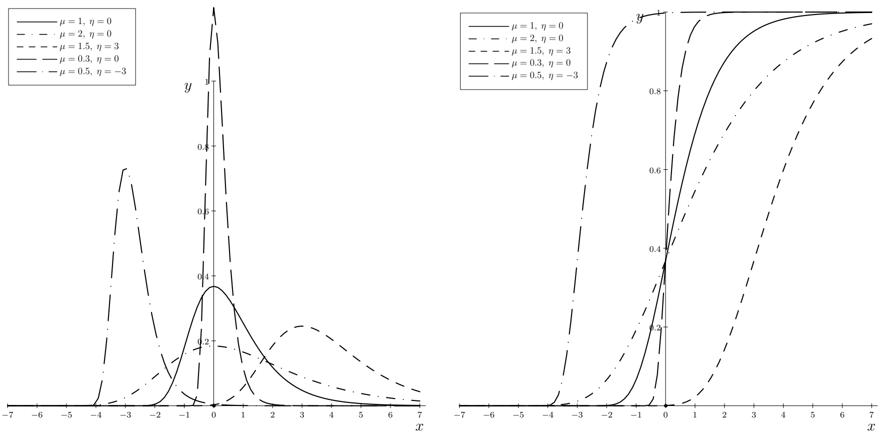
in which μ is a positive scale parameter and η is a location parameter. The mean of this distribution is given as $\eta + \gamma\mu$ with γ being Euler's constant. The variance is $\pi^2/6\mu^2$. The density and cumulative distribution for different values of μ and η are given in figure 4.1.

The Gumbel distribution is chosen because it leads to a structural solution of the probability function in equation (4.2). The MNL model is expressed as:

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum\limits_{j \in C_n} (e^{\mu V_{jn}})} \tag{4.5}$$

where:

$P_n(i)$ = Probability P of choosing alternative i by individual n
C_n = Choice set of individual n
V_{in} = Systematic utility of alternative i for individual n
μ = Positive scale parameter



(a) *PDF* **(b)** *CDF*

Figure 4.1: *Plots of the density and cumulative distribution function of the Extreme Value Type 1 distribution for different values of μ and η*

In case of a linear-in-parameter utility function V_{in} the probability of choosing alternative i is given by:

$$P_n(i) = \frac{e^{\mu \beta_i x_{in}}}{\sum\limits_{j \epsilon C_n} \left(e^{\mu \beta_j x_{jn}}\right)} \tag{4.6}$$

which can be estimated using maximum likelihood. As is clear by this equation (4.6), it is not possible to distinguish the parameter μ from the overall scale of the β parameters, as the combination $\mu\beta$ is estimated. As a convention, the scale parameter is commonly assumed to be as $\mu = 1$, which assumes the variance of all random or unobserved parameters $\varepsilon = \pi^2/6$. This does not mean the scale parameter μ is of no importance. Contrary, by assuming it is equal for all unknown parameters ε, we assume homoscedasticity, indicating that the variance is equal for all cases, either alternative, choice moment or individual.

If we assume multiple segments in the dataset with different variance, it is possible to estimate a scale parameter for all minus one of the segments (Train, 2009). This scale parameter then indicates the relative variance of the other segments to a segment that is chosen to be normalized. Assume a linear-in-parameters utility function for two segments which uses the same alternatives and attributes but has a different variance of the unobserved term:

$$\begin{aligned} U^1_{ni} &= \beta^1 x_{ni} + \varepsilon^1_{ni} \\ U^2_{ni} &= \beta^2 x_{ni} + \varepsilon^2_{ni} \end{aligned}$$

Herein the variance of segment 1 ε^1_{ni} is different from the variance of segment 2 ε^2_{ni}. If the ratio of the variance is given by $r = \frac{Var(\varepsilon^2_{n1})}{Var(\varepsilon^1_{n1})} = \frac{\mu_2^2}{\mu_1^2}$ it is possible to rewrite both equations assuming equal taste variations for both segments ($\beta^1 = \beta^2$):

$$
\begin{aligned}
U_{ni}^1 =& \beta x_{ni} + \varepsilon_{ni}^1 \\
U_{ni}^2 =& \sqrt{r}\beta x_{ni} + \varepsilon_{ni}^1
\end{aligned}
$$

This shows that the variance of segment 1 ε_{ni}^1 is now the same for all segments. The parameter is estimated together with the other coefficients in the model. The estimated value of r indicates the variance of the random factors in segment 2 relative to that in segment 1, in case the scale parameter of the first segment is normalized to one.

4.2.2 Independent and identically distributed errors

The MNL model is often the basis of the discrete choice models, as it is relative easily estimated using available software packages such as Biogeme (Bierlaire, 2008), it is flexible in the attributes to be used and often results in taste parameters which can easily be explained. However, the MNL model also has an important property which might act as a limitation: Independence of Irrelevant Alternatives (IIA). This property states that the ratio of the probabilities of any two alternatives is unaffected by the systematic utilities of any other alternatives. This property can be seen in equation (4.7), ignoring the scale parameter μ for alternative i and k:

$$
\frac{P_{ni}}{P_{nk}} = \frac{e^{V_{ni}} / \sum_j e^{V_{nj}}}{e^{V_{nk}} / \sum_j e^{V_{nj}}} = \frac{e^{V_{ni}}}{e^{V_{nk}}} = e^{V_{ni} - V_{nk}} \tag{4.7}
$$

No matter how simple this property may seem at first, the usage of it can lead to obviously wrong results in an estimated model. A very famous example in this case is the "Red Bus-Blue Bus" problem, in which a standard multinomial logit model would lead to counter-intuitive results. Assume we have only two modes available to choose from: car and a red bus, where both modes have a market share of ½. When a new mode is introduced, namely the blue bus, which has exactly the same attributes as the red bus, except for the colour, one would expect the market share of the red and blue bus together to be equal to ½. However, the multinomial logit model will give a different market share; ⅓ for all three modes, because of the irrelevance of other alternatives. The ratio of car and red bus has to be equal in both cases, which can only occur if the market share of cars and red buses is ⅓. This is opposite of expectations, where the market share of car stays the same (½) and the market share of either blue or red bus is ¼. It is also possible to see that this property will lead to erroneous results in a route choice with overlap in the routes (Ben-Akiva & Bierlaire, 2003).

The IIA property of the MNL model can thus be interpreted as a failure to account for similarities of the alternatives. The assumptions of the MNL model (independently distributed random terms), mean that unobserved correlations (or similarities) are not included in the model. If this assumption is not justified, the estimated parameters are most likely biased. The IIA property also has its advantages, as it allows a researcher to estimate a subset of alternatives, as these are independent of the alternatives excluded.

Another possible violation of the assumption of independent and identically distributed errors exist: repeated choices. Often, the choices of a decision maker are sampled at multiple points in time. This can either be in real life, for example the choice for an exit at each junction or for a route at each day, or in a hypothetical environment, where a decision maker is asked multiple times to choose between different alternatives. Data that captures multiple observations in time from one decision makers is called panel data. If the unobserved terms in the utility function are independent over these repeated choices, the assumptions for the MNL model hold and can be applied as if it is cross-sectional data. However, one would expect that some of the unobserved terms are dependent on the point in time. This violates the assumption of independent errors, and introduces correlation over time. This means the utility function includes a new index t for time:

$$U_{int} = V_{int} + \varepsilon_{int} \tag{4.8}$$

In general, three ways are available to relax the i.i.d. assumption of the MNL model:

- allowing the random terms to be identically distributed while being non-independent (being correlated);
- allowing the random terms to be non-identically distributed (having different variances) but maintaining the assumption of independence;
- allowing the random terms to be both non-identical and non-independently distributed.

The first way of relaxing the i.i.d. assumption allows for a correlation of the unobserved terms, while maintaining the same distribution. Using the Generalised Extreme Value (GEV) type models, it is possible to determine this correlation between alternatives of a certain segment of which the nested logit model is the most common. The nested logit model will be presented in section 4.2.3.

The second way introduces an unconstrained variance of the random terms, which essentially means that the alternatives in a choice set can have different variances. The heteroscedastic extreme value (HEV) logit model is based on this approach and is presented in section 4.2.4. The third way relates to using a complex covariance matrix which complicates the estimation as well as interpretation of the results. However, when using the normal distribution for the unobserved terms it is possible to estimate these values. Also the mixed logit model allows to deal with both non-identical and non-independently distributed error terms. The mixed logit model will be discussed in section 4.2.5.

In all cases, the researcher is interested in accounting for heterogeneity and heteroscedasticity. Heterogeneity refers to the variation across individuals in the propensity to select different alternatives and responsiveness to independent variables. This can be seen as a sort of taste variation, correlation or covariance of alternatives and their attributes. Heteroscedasticity refers to the variance of individuals or alternatives.

4.2.3 Nested logit

Nested logit was introduced by several authors independently (Daly & Zachary, 1978; McFadden, 1978; Williams, 1977). In nested logit, all the alternatives j are partitioned into a number K of non-overlapping subsets, or nests B_k. In case of a two-level choice model with a generic choice set (G) and an elemental choice set (M) the utility function can be given as (see for a simple example figure 4.2):

$$U_{gm} = U_g + U_{m|g}, m \in M, g \in G \tag{4.9}$$

$$U_{gm} = V_g + V_{m|g} + \mu_g + \varepsilon_{m|g}, m = 1, \ldots, M_g, g = 1, \ldots, G \tag{4.10}$$

The utility function is thus divided into a part for the utility of the generic choice set, and conditional on the choice of the generic branch and a part for the utility of the elemental choice set. The joint probability of choosing alternative gm is then defined as a conditional probability of choosing alternative m, given nest g, and a marginal probability of choosing nest g:

$$P_{gm} = P_{m|g} P_g \tag{4.11}$$

$$P_{gm} = \frac{e^{\lambda_g V_g + \frac{\lambda_g}{\lambda_m} \log \sum\limits_{m\prime \in Mg} e^{\lambda_m V_{m\prime|g}}}}{\sum\limits_{g\prime \in G} e^{\lambda_{g\prime} V_{g\prime}} + \frac{\lambda_{g\prime}}{\lambda_{m\prime}} \log \sum\limits_{m\prime \in Mg} e^{\lambda_{m\prime} V_{m\prime|g}}} \cdot \frac{e^{\lambda_m V_{m|g}}}{\sum\limits_{m\prime \in Mg} e^{\lambda_{m\prime} V_{m\prime|g}}} \tag{4.12}$$

in which:

V_g = Utility of generic alternative g
$V_{m\prime|g}$ = Utility of elemental alternative m given generic alternative g
λ_g = Scale parameter of generic alternative g
λ_m = Scale parameter of elemental alternative m

When the ratio of the scale parameters is equal to one, the nested logit model collapses into the multinomial logit model. This also shows in the calculation of the correlation between two nests, which is based on:

$$Corr = 1 - \left(\frac{\lambda_g}{\lambda_m}\right)^2 \tag{4.13}$$

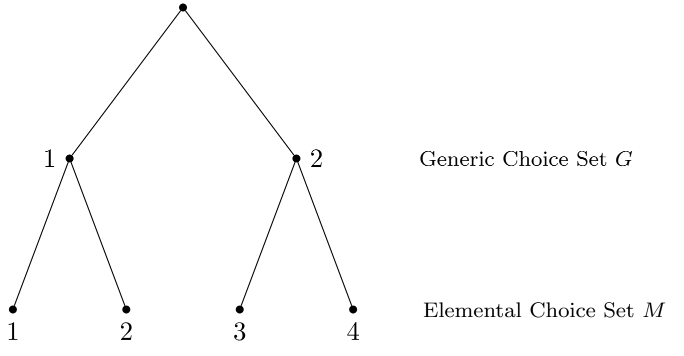


Figure 4.2: *Generic nest structure for Nested Logit models*

When the correlation is zero, there is no common element in the utility functions of the alternatives, which thus reduces to the multinomial logit model. For the nested logit model to be consistent with random utility maximization, it is necessary that the ratio of the scale parameters is between one and zero. This means that the inequality $\lambda_g \leq \lambda_m$ must hold true.

In order for the model to be estimated, it is necessary to normalize one of the scale parameters to one. There are two ways of doing this (Hensher & Greene, 2002b): from the bottom, called RU1, or from the top, called RU2. In RU1, the scale parameter of the elemental alternatives are set to one, while in RU2, the scale parameters of the generic alternatives is set to one. Normally both approaches will give the same results. It is important to note that the scale parameters within a nest or branch have to be equal (so λ_m for $m = 1, 2$ is equal). More information about this can be found in several papers, of which Hensher & Greene (2002a) is a good starting point. When estimating a nested logit model with degenerate branches which means that a nest contains only one elemental alternative, the scale parameter for this nest equals out. In other words, the utility function of this alternative is not specified at the elemental alternative level, but at the level of generic alternatives.

4.2.4 Heteroscedastic extreme value logit

The MNL approach is based on the assumption that the unobserved terms are independently and identically distributed. This limits the variances to be equal for all alternatives. However, this assumption might be invalid if the variance of an unobserved (or not included) parameter differs over alternatives. For example, if the unobserved parameter road type varies greatly for a left exit, while it is rather stable for the straight ahead exit, the random terms in equation (4.1) will have different variances for different alternatives.

In case of the heteroscedastic extreme value logit (Allenby & Ginter, 1995; Bhat, 1995; Hensher, 1997), the scale parameter is defined to be dependent on the alternative (setting μ as μ_i in equation (4.3)), with the normalization that the error terms of one of the alternatives have a scale parameter of one for identification. This leads to the probability of choosing alternative i as:

$$
\begin{aligned}
P_i =& Pr\left(U_i > U_j\right) \forall j \in C, i \neq j \\
=& Pr\left(\varepsilon_j \leq V_i - V_j + \varepsilon_i\right) \forall j \in C, i \neq j \\
=& \int_{\varepsilon_i=-\infty}^{\varepsilon_i=\infty} \prod_{j \in C, j \neq i} \Lambda\left[\frac{V_i - V_j + \varepsilon_i}{\mu_j}\right] \frac{1}{\mu_i} \lambda\left(\frac{\varepsilon_i}{\mu_i}\right) \mathrm{d}\varepsilon_i
\end{aligned}
$$

Herein Λ is the cumulative distribution function and λ the probability distribution function of the type I extreme value distribution. When substituting $w = \frac{\varepsilon_i}{\mu_i}$ this results in:

$$
P_i = \int_{\varepsilon_i=-\infty}^{\varepsilon_i=\infty} \prod_{j \in C, j \neq i} \Lambda\left[\frac{V_i - V_j + \mu_i w}{\mu_j}\right] \lambda(w) \mathrm{d}w \tag{4.14}
$$

The HEV model thus allows for different scale parameters across alternatives. As equation (4.14) shows, the larger the scale of alternative i, the smaller the difference between alternative i and j will be. A change in the observed utility of alternative j will thus have less influence on the probability of choosing alternative i, if the scale parameter of alternative i is large (and hence, the variance of alternative i is small). The cross elasticity between alternative i an j is thus dependent on the scale parameters. The IIA property no longer holds for this model.

4.2.5 Mixed logit

Mixed logit models have been known for many years and saw the first applications in 1980 (Boyd & Mellman, 1980; Cardell & Dunbar, 1980). Only in the more recent years, the mixed logit model has been widely applied as the necessary simulation is computationally much faster nowadays. For a detailed analysis of the mixed logit model we refer to McFadden & Train (2000); Hensher & Greene (2003); Train (2009). Mixed logit does not, contrary to multinomial logit, have the IIA property. The ratio of two alternatives depends on all the data, whereas in multinomial logit, only the data of the two alternatives is necessary.

The mixed logit model assumes that the observed portion of the utility function depends on the parameters β, for which a distribution is assumed with a density $f(\beta \mid \theta)$ where the parameters of the distribution are given by θ. For a given value of β, the (conditional) probability of the choice of alternative i by individual n is given in equation (4.15).

$$L_{ni}(\beta \mid \theta) = \frac{e^{V_{ni}(\beta)}}{\sum\limits_{j=1}^{J} e^{V_{nj}(\beta)}} \tag{4.15}$$

Integration over all values of β leads to the unconditional choice probability as in equation (4.16).

$$P_{ni} = \int L_{ni}(\beta) f(\beta \mid \theta)\, d\beta \tag{4.16}$$

The name of mixed logit is explained as the choice probability is a mixture of logit with a mixing distribution f. Mixed logit assumes a general distribution for β; it can take many different forms such as normal, lognormal, uniform or triangular. Depending on this mixing distribution different substitution patterns may be obtained for the alternatives. The first is known as random parameter or random coefficients. In this approach, each element of the vector β is associated with an attribute of the alternative i distributed with a mean and standard deviation (or any other definition of a distribution). The second approach is known as error components, which treats the unobserved terms as two separate error components.

In case of random parameters, the utility of person n for alternative i is described as:

$$U_{ni} = \beta_n^{'} x_{ni} + \varepsilon_{ni} \tag{4.17}$$

where $\beta_n^{'}$ is a vector of coefficients of the observed variables x_{ni}, with a density function $f(\beta \mid \theta)$. This is the same specification as for the normal logit model, apart from the β that varies over individuals, instead of being fixed.

In case of error components, utility is defined as:

$$U_{ni} = \beta_i^{'} x_{ni} + \eta_n^{'} z_{ni} + \varepsilon_{ni} \tag{4.18}$$

in which both x_{ni} and z_{ni} are vectors of observed variables related to alternative i. β_i is a vector of fixed coefficients related to each alternative and η_n is a vector of random terms with zero mean. ε_{ni} is i.i.d. extreme value type 1 distributed. The terms in z_{ni} are named error components that together with ε_{ni} described the random portion of the utility. If the error components are nonzero, the alternatives are correlated. Using error components it is possible to create various correlation patterns, by selecting the appropriate variables. Both approaches are formally the same, as both x_{ni} and z_{ni} can overlap.

The specification of the mixed logit model is easily generalized to incorporate repeated choices over individuals, such as in the design for this research. The utility function that describes these choice situations t is:

$$U_{nit} = \beta_n^{'} x_{nit} + \varepsilon_{nit} \tag{4.19}$$

where ε_{nit} now is i.i.d. extreme value over individual, alternatives and choice situations. The conditional probability of making a sequence of choices over the alternatives is the product of logit formulas:

$$L_{ni}(\beta) = \prod_{t=1}^{T} \left[\frac{e^{\beta_n^{'} x_{ni_t t}}}{\sum_j (e^{\beta_n^{'} x_{njt}}} \right] \tag{4.20}$$

where $\imath = \{i_1, \ldots, i_T\}$. The unconditional probability is the integral of equation (4.20):

$$P_{ni} = \int L_{ni}(\beta) f(\beta)\, d\beta \tag{4.21}$$

The main difference between the "normal" mixed logit model, and the mixed logit model for panel data is that the integral is taken over a product of logit formulas.

An important issue for mixed logit is that an exact computation is not possible, as was the case for multinomial and nested logit. Instead, simulation is necessary to approximate the integral in the unconditional probability (Train, 2009). Using simulation, a draw of β is taken from the density $f(\beta \mid \theta)$, with which the conditional probability is calculated. By taking many draws, and averaging the results, a simulated unconditional probability is calculated, which is an unbiased estimator of the true probability. Several methods are available to generate the draws, of which the Halton sequences provide the most "intelligent result". Also the number of draws is important. A recommended number of draws is 1000. Next, depending on the approach chosen and the attributes used for taste variation, identification issues are important, as stressed by Walker (2001); Walker et al. (2007). We refer to these papers for further information of these issues.

4.2.6 Test methods for discrete choice models

The first test that should be applied to the results of a random utility model is an informal test. It comes down to merely looking at the sign of the estimated parameter and the size of it compared to others. If these are in accordance with what could be expected, the model provides at least reasonable results. However, more statistical approaches are necessary to provide more insight into the results.

A number of statistics are available to provide insight into the significance of the results. A first test is the asymptotic t-test, to check whether a particular estimate is significantly different from a constant, which normally is zero. The critical value for a significance level of 5% would be ± 1.96.

Another basic test is to see if the estimated parameters jointly differ from zero. The likelihood ratio test can be used for this hypothesis. This test is defined as:

$$-2(\mathcal{L}' - \mathcal{L}^*) \tag{4.22}$$

in which:

$\mathcal{L}'$ = Null Log Likelihood for the estimated model with all β's and alternative specific constants being zero

$\mathcal{L}^*$ = Final Log Likelihood for the estimated model with optimal β's and alternative specific constants

The resulting value of this test is χ^2 distributed with degrees of freedom equal to the number of estimated parameters. If the test value exceeds the critical value of the χ^2 distribution, the null hypothesis of all parameters being equal to zero is rejected. This test can also be applied to check if two estimated models differ significantly over all parameters. In that case the Null Log Likelihood is replaced by the Final Log Likelihood of the other estimated model. The absolute difference in number of estimated parameters then defines the degree of freedom for the test, and enables one to look up the critical value.

To test for the IIA property of the MNL model, it is possible to estimate two models: one for the complete choice set and one for a subset. If the model is correctly specified, the estimated coefficient should be consistent over both models. Denote the estimated coefficients for the full model to be β_C and for the subset $\beta_{\hat{C}}$. Several test statistics are developed (Ben-Akiva & Lerman, 1985; Cheng & Long, 2007) for the null hypothesis that $\beta_C = \beta_{\hat{C}}$, for instance:

$$(\beta_{\hat{C}} - \beta_C)(\sum_{\beta_{\hat{C}}} - \sum_{\beta_C})^{-1}(\beta_{\hat{C}} - \beta_C) \tag{4.23}$$

which is asymptotically χ^2 distributed with $\tilde{K}$ degrees of freedom. $\tilde{K}$is the number of elements in the subset of coefficients that is identifiable using the subset of the complete choice set. Another test statistic is:

$$\frac{1}{1 - N_1/(\alpha N)}\{-2[\mathcal{L}_{\hat{\mathcal{C}}}(\beta_{\mathcal{C}}) - \mathcal{L}_{\hat{\mathcal{C}}}(\beta_{\hat{\mathcal{C}}})]\} \tag{4.24}$$

in which N is the number of observations in the complete choice set and N_1 is the number of observations in the restricted choice set and $\alpha \geq 1$ is a scalar. $\mathcal{L}_{\hat{C}}$ is the log likelihood value calculated on the estimation sample using the subset of the choice set. Again, this test statistic is asymptotically χ^2 distributed with $\tilde{K}$ degrees of freedom.

Goodness-of-Fit measures are also available for estimated models. The ρ^2 and pseudo-ρ^2 are available for this purpose, which are similar to R^2, which is used in regression analysis. ρ^2 is given by:

$$\rho^2 = 1 - \frac{\ell^*}{\ell^0} \tag{4.25}$$

and the pseudo-ρ^2 is given by (in which K is the degrees of freedom):

$$\rho^2 = 1 - \frac{\ell^* - K}{\ell^0} \tag{4.26}$$

If the values of ρ^2 are between 0.2 and 0.4, the model is indicated to have a very good fit (Louviere et al., 2000), which is more or less equivalent to 0.7 and 0.9 of R^2.

A last test relates to the validity of the models that are estimated. This informal test compares the observed market shares for the alternatives in the choice set with predicted market shares based on the estimated models. This implies an aggregation of the resulting choice probabilities. However, as often the only the same dataset is available for comparison of the market shares, the prediction is likely to be equal to the observation. By reducing the dataset which is used for estimation (for example by randomly removing a choice situation), it is possible to use these removed choice situations for prediction. As the actual choice is observed, the predicted and observed market shares can be compared. The closer the prediction, the better the model.

4.2.7 Validity, limitations and advantages

Regarding the use of discrete choice models it is important to discuss the validity of the results and the approach. Validity in a discrete choice model can be either internal or external validity. Internal validity concerns the expected signs of the coefficients. The external validity concerns the predictive value of the results. It can be separated into two items: (a) concurrent validity in stated preference and revealed preference and (b) predictive validity. (Carson et al., 1994; Beaton et al., 1998). The concurrent validity aims at determining whether there is a significant difference between resulting choice probabilities which stem from choice models based on either stated preference data or revealed preference data. The predictive validity focuses on prediction; either a stated preference or revealed preference model is used to predict choice probabilities, based on certain policies. Testing for predictive validity thus occurs only after an actual change has happened, which can be incorporated in a choice model. The difference in both types of validity is obvious; a model that shows concurrent validity (and thus has much explanatory power in its estimates) does not necessarily have to have the ability to predict future choices. Predictive validity is therefore the most restrictive type. In many previous studies regarding validity of discrete choice models, the validity has been investigated, for transport and traffic applications, as well as many other areas. In these studies, results are often found to be valid, or have reasonably low prediction errors (Carson et al., 1994; Beaton et al., 1998).

The approach of the discrete choice models presented in this section is based on the choice for the alternative with the maximum utility. This assumption is often debated, as humans are not likely to maximize the available options. Several other approaches are available which for example are based on minimizing regret or different valuation of gains and losses compared to a status quo (see for overviews Chorus et al. (2009); Van de Kaa (2010)). These approaches are different to the maximization principle, by accommodating several behavioural aspects such as habit, regret or reference states.

Next to the validity of the underlying approach, there are some other limitations of random utility maximization as described in this section, similar to those described in Bogers (2009). First, in case of habit, there choice is automatic and the choice maker is not aiming for maximization. Second, varying emotions or risk approaches are difficult to incorporate directly, and are often only possible to determine in the unobserved (random) terms. Last, each attribute is of a compensatory nature. A bad value for an attribute can be compensated for by a good value for another attribute, whereas in reality, the bad value of an attribute might cause a choice maker to disregard that alternative as a whole.

However, the approach also has its advantages for modelling. First, the resulting estimation of the utility functions provide information about the (relative) importance of attributes for decision making. For example, a value of costs which is higher than zero and reliability, indicates that costs are of importance in decision making, and more important than reliability. The resulting models can be compared using the likelihood ratio test and on the pseudo ρ^2, which enables a researcher to compare different specifications and select those that provide the best result. Last, the theory behind utility maximization is well developed and continues to receive attention, for its simplicity in modelling. Many models have been developed and applied (of which a few have been presented here) and software for the estimation of these models is readily available.

Even though, the approach of maximizing utility can be questioned on its validity, the results of these models provide significant input about the choice behaviour of humans. However, as with all models, the models are used as a means to an end, in which the end is an indication of the importance and relations in the choice making process.

4.3 Route choice modelling

Routes are a specific instance of the alternatives for a discrete choice model. A route is defined by Bovy & Stern (1990) as "a chain of consecutive road segments connected by nodes. Such chains connect trip origins and destinations." A route choice is in principle the same as a discrete choice, as the alternatives meet the requirements for such a choice set:

- the set of routes is exhaustive, meaning that it includes all alternatives available (for an individual);
- the routes are mutually exclusive, as a person can only choose one route in a single choice situation;
- the set of route contains a finite number of alternatives (when excluding loops).

However, route choice as a discrete choice problem has its specific characteristics as given by Ben-Akiva & Bierlaire (2003). First, the universal choice set is very large, as the road network can provide a very large number of routes. Second, a decision maker is unlikely to consider all these alternatives, but will use a subset, which is often unknown to an observer. Third, the overlap in the alternatives means the alternatives are correlated, violating the assumptions of the MNL model.

Route choice is a dynamic choice, as it can happen at different time instances and different choice situations. A route choice can be described for:

- day-to-day route choices;
- within day route choices;
- pre-trip route choices;
- en-route route choices.

It is possible to even determine a fifth situation, related to en-route route choices, which determines the choice to stay on the current route or switch to another alternative route, which can be thought of as a binary choice.

Discrete choice modelling is not the only approach available for route choice modelling. Other approaches such as intelligent agents (Bazzan et al., 1999; Dia, 2002; Dia & Panwai, 2007), fuzzy logic (Lotan & Koutsopoulos, 1993; Lotan, 1997; Ridwan, 2003; Quattrone & Vitetta, 2011) and neural networks (Yang et al., 1993) have been applied. We do not further discuss these approaches but focus on using discrete choice modelling as discrete choice modelling enables us to estimate various models and assumptions which can directly be related to the used approach. Next, the estimated coefficients are of importance in determining effects of, for example, traffic information on route choice behaviour.

Frejinger (2008) mentions three main steps before a route choice model can be estimated (see figure 4.3). First of all, it is necessary to estimate a model based on data. In the case of route choice modelling, this concerns trips undertaken by drivers, for which a route was chosen. Second, the alternatives that are considered by the driver have to be included in the choice set. These alternatives are generally unknown to the observer. As the amount of possible alternatives is very large in a realistic road network, this poses a difficult task for a observer. Third, before

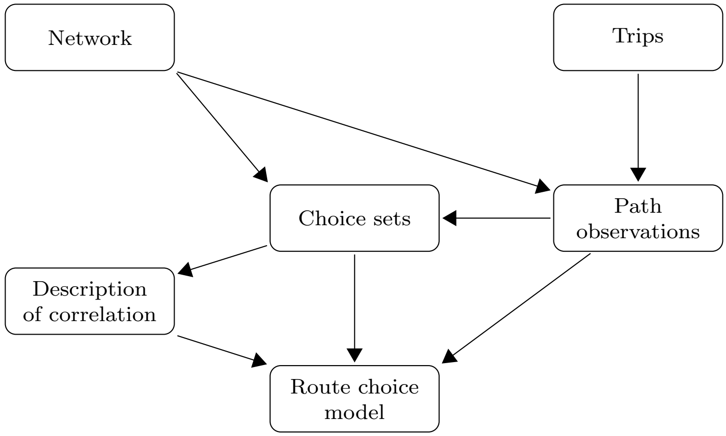


Figure 4.3: *Route choice modelling overview (Frejinger, 2008)*

estimating the choice model, it is necessary to account for the overlap of the routes in the choice set. The routes in the choice set are likely to share some road segments, and thus are correlated. As mentioned in section 4.2.2 this violates the assumptions of the MNL model. All these subjects have been discussed numerous times in transport and related literature. This section will not discuss all the available theories, but focus on the route choice related to this research. We refer to Prato (2009) for an overview.

4.3.1 Trips and route observations

A traveller makes numerous trips in his daily life. For each trip, multiple choices are made, regarding activities, destinations, modalities and routes, sequentially or simultaneously. Of interest for this thesis are trips made with a car, especially in case of non-recurrent conditions. The routes chosen in these trips need to be observed.

Depending on the type of route choice model to be estimated, the observation needs to be adapted. In case of day-to-day route choice, multiple observations of one driver are necessary, whereas for a pre-trip route choice, one could suffice with just one observation of each driver. For modelling en-route choices, at least a full trip is necessary. In all of the previous cases, an origin and destination are necessary to define the start and end of the route. For observing route switching decisions, one needs more than just observing the chosen route. It is necessary to have additional information on the currently selected route, in order to know if the driver changed his route.

Observations of actual routes can come from multiple sources. Surveys (using the internet, phone, paper or interviews) are a conventional way of asking travellers for the routes they chose. This often results in routes that are described at an overall level, as it is a difficult and time consuming task to describe each individual route in great detail. Next, it is often difficult or even impossible to acquire all the necessary information for these route observations, which is important in non-recurrent situations. Another source of observations uses GPS to track travellers, which can be combined with a survey. GPS is a passive monitoring technique, which stores the location where a traveller has been, with a certain time or distance interval between the locations. Based on these points, a route can be reconstructed, which is however not a straightforward process. The difficulties and an approach to reconstruct routes from GPS data is given in Schüssler (2010). A third method would imply the tracking of license plates at multiple locations in the road network. However, these observations do not include an origin and destination, which makes it difficult to fully model the route choices. Outside of real world observations it is also possible to observe hypothetical route choices, in surveys or simulations.

A route is difficult to give an unambiguous observable label (Bovy, 2007). This implies that the route which is selected, cannot be described other than using its generic attributes, together with the set of links. This is an important distinction with many other applications of discrete choice modelling, in which the choice (and the alternatives under consideration) are given explicit and observable labels, which enables an analyst to the describe the alternatives with their specific (labelled) attributes. In case of a route switching decision, this generic approach is avoided if the choice is between staying on the current (pre-trip) selected route, or switch to a different route. In case of selecting a specific link at each decision point (junction) in the road network, it would also allow the analyst to describe the observations in terms of left, right or straight ahead choices, and thus create labelled alternatives.

A full overview of the different sources and types of route choice data will be given in section 6.2. For the route choice model in this study we are interested in the en-route decisions at individual junctions. As this research concerns non-recurrent traffic situations in an urban area, drivers are able to select routes which are off the main routes. Together with ATIS, a driver is expected to reconsider route choice at each decision point in the road network. The type of observations needed are thus the choice for an exit at each decision point.

4.3.2 Choice set generation

The choice set is the set of alternatives which is used by the researcher to estimate a route choice model, and is necessary in case of an application of such a route choice model. As mentioned, in the case of route choices, the choice set of a realistic road network is fairly large, as a large amount of routes can be constructed. A choice set can be used for three separate purposes (Prato, 2009):

1. analyzing the travel alternatives to determine their availability, number, characteristics, variety and composition;
2. estimation of disaggregate demand models to uncover behavioral parameters of utility functions at the individual level, using observations of individual route choices;
3. prediction of choice probabilities to determine route and link flow levels in networks, using route choice models with estimated parameters.

In this research we are interested in the second purpose; the estimation of a disaggregate model, which allows one to estimate the behavioural parameters of

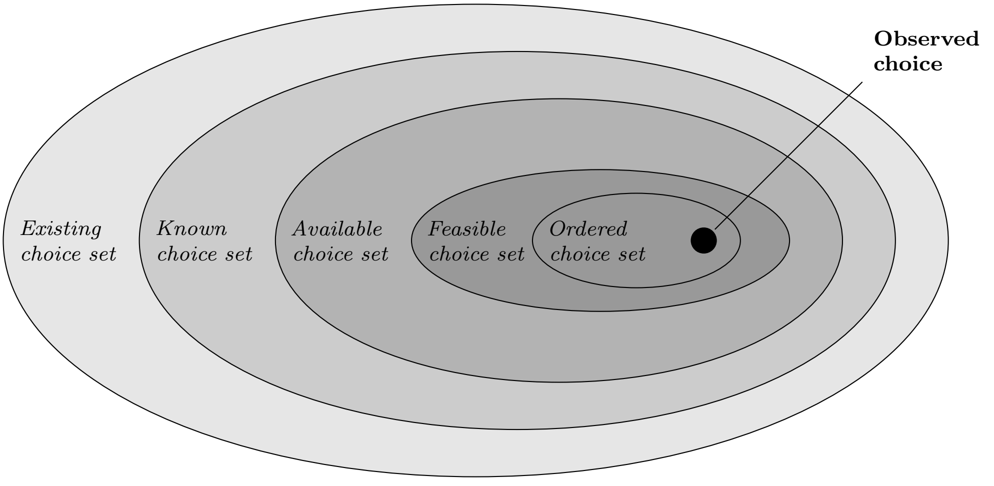


Figure 4.4: *Simplified choice set overview Bovy & Stern (1990)*

the utility function of routes.

Regardless of the purpose for which a choice set is "created", choice sets with respect to route choice modelling have a number of specific aspects (Bovy, 2007):

- the number of available alternatives (the universal or master choice set) in real road networks is very large;
- the subset of feasible and attractive or considered routes is also fairly large and difficult or even impossible for an analyst to acquire;
- the feasible, attractive or considered choice set is likely to be dynamic, as the interaction of demand and supply change the attributes of the routes in these choice sets;
- depending on the choice process (sequential, simultaneous or strategic), the choice set has a different size and compositions and needs to be recreated at each decision point.

An excellent discussion of the concepts related to the different choice sets and their composition is given in Hoogendoorn-Lanser (2005). An overview of the different choice sets is given in figure 4.4, based on the conceptual framework of a route choice process given in Bovy & Stern (1990).

Based on the transport system or road network, for each origin and destination all routes together form the existing route alternatives or the existing choice set, which is the universal or master choice set. Travellers only have a limited knowledge of the available routes. This knowledge is based on their previous experiences and that of others (by asking) and other information sources such as maps and navigation devices. Together with a different perception of these alternatives, the set of known alternatives is specific for each individual. Based on the constraints which are set by the travellers themselves and their preferences, the set of known alternatives is reduced to the available set of alternatives. This set is further reduced to the feasible alternatives by eliminating those alternatives which do not match the conditions of that specific trip to be made, such as time or monetary feasibility conditions. Eventually, not all the feasible alternatives are considered in the choice process, as the number of alternatives which is often subject to the decision rule is relatively small. Around 5 alternatives are expected to be a common size for the set of ordered alternatives. As pointed out by Hoogendoorn-Lanser (2005), it is important to distinguish the view point of the choice set, as there is a possible disagreement between the actual individual choice set, the reported individual choice set and the generated individual choice set (by the researcher). Bovy (2009) concludes that based on the processes involved in route choices we may expect strong individual differences in behaviour caused by differences in knowledge, perceptions and preferences for the alternatives and thus for the actual choice. The separation of the choice set formation and the choice is often advantageous as the process are different and it allows for a analysis of different impact studies on both the choice and the choice set formation (for example giving route advice, which influence the actual route choice, but also the knowledge of alternatives).

Regarding the use of a choice set to estimate a route choice model, different results are found. Bovy (2009) argues, that it is not necessary to include all relevant alternatives in the choice set, as satisfactory results for the estimated coefficients in utility functions are obtained even for small well-sampled choice sets, based on correction of the model for the sample. We refer to Ben-Akiva & Lerman (1985) for more details on sampling correction in case of discrete choice. However, others (Bekhor et al., 2006; Prato & Bekhor, 2007; Bliemer & Bovy, 2008) argue that the size of the choice set and the alternatives it contains affect the estimated coefficients and choice probabilities for several model specifications. Ideally, all relevant and no irrelevant alternatives should be included in the choice set.

In order to generate the necessary choice set with only relevant alternatives, numerous approaches are available. Prato (2009) and Bovy (2009) provide a detailed overview of the current state of generation approaches and we refer the reader to these papers for more detail. They both classify the generation approaches as follows:

- probabilistic;
- constrained enumeration methods;
- deterministic shortest path methods;
- stochastic shortest path methods.

Probabilistic methods attach a probability to each route. For example Frejinger (2007); Frejinger et al. (2009) calculate a probability for each link in the network. The probability of a link depends on its distance from the shortest route according to a certain cost function, which means that all links on the shortest route have a probability of one and all other links a probability between one and zero. Using a random walk, the probability of a route between origin and destination is found, which is used to correct the sampling probability in case of estimation. The constrained enumeration methods generate an exhaustive universal set given a set of constraints, by building a full connection tree between the origin and destination for which the additional links are added in case these satisfy the constraints, such as cognitive, perceptual and behavioural. A sequence of links enters the choice set if it connects the origin and destination and satisfies the conditions. As such, this method relies on the behavioral assumptions of the travellers, which are expected to be of importance by the researcher.

The largest group of generation methods is based on a repeated shortest route search. The computation of different paths is based on a modification of one or more input variables such as search criteria or route constraints. The most straight forward approach is to calculate K shortest paths. A different approach assumes drivers have different objectives. These objectives may correspond with different routes and each of these routes can be labelled according to a different objective function. In all cases, the routes are deterministic which means they are the same each time they are generated in the same network. In case of stochastic shortest path methods, the path finding applies distributions to the cost functions of the routes, making the routes that are found to be random. This is based on the assumption that drivers perceive the route attributes such as travel time with an error.

In all cases, these generation methods are important in case of a realistic road network where the number of alternatives is large. In a stated preference research, where the researcher is able to specifically design and present the alternatives, it is expected to be of less importance, even though Bovy (2009) mentions the same approach can be applied. This largely depends on the design of the experiment (by the number of presentation of the alternatives). Another important aspect in the generation of these methods is that they rely on a route between the origin and destination. In case of a sequence of route choices (en-route), at each decision point, a new choice set would be needed. This implies a huge impact on computation effort in a realistic road network.

To avoid this problem one can resort to using a route switching model, where a generation method for the choice set is not necessary, as the choice is a binary one. Another possibility is to use the available links at a decision point as the actual choices. This resorts to (on average) a maximum of three alternatives (excluding a u-turn). Assuming a driver selects his/her route at each of these decision points, it is possible to only use the routes starting at these alternatives. This greatly reduces the computational effort, and one only needs to apply a search algorithm to find the "optimal" route starting at that link towards the destination. This approach is used in this research.

4.3.3 Route choice model

Once the alternatives are available, the modelling of the choice between the alternatives is of interest. However, as depicted in figure 4.3, in case of route choice modelling, there is bound to be overlap in the routes of the choice set. As overlap in the links of the various routes violates the assumption of independence of unobserved terms for alternatives which is used for the MNL model, this model is theoretically not suitable for route choice situations. The same applies to Nested Logit models, as these models assume an alternative to belong to one nest exclusively, whereas a route often share links with many other routes. This overlap problem is depicted in figure 4.5.

Three routes are available to get from A to C. If all routes are assume to have equal utility (based on for example equal length or travel time), the probability of

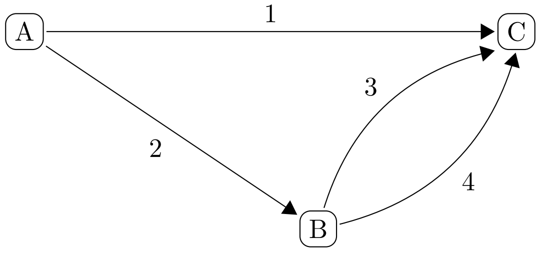


Figure 4.5: *Overlapping routes*

choosing each route is equal to 1/3 based on the MNL model. This is regardless of the attributes of the common link 2. If this link is large (in length or travel time) compared to the links 3 and 4, it is expected that the two routes using link 2 in total will have a more or less equal probability compared to the route using link 1. However, the MNL model will not predict such a result.

Describing the correlation of the alternatives in the choice set is possible in many different ways. The first group is based on correcting the MNL model. The second is based on the Generalized Extreme Value models and the third is based on mixed logit.

In case of the MNL model, several correction factors are proposed, which all describe the overlap of the routes in the choice set. The assumption of the choice set is that it contains full routes towards the destination. The three factors presented here are the C-Logit model, the Path Size Logit model and the Path Size Correction Logit model.

The C-Logit model was proposed by Cascetta et al. (1996); Cascetta & Papola (2001), in which a correction factor is introduced. This commonality factor (CF) is used to correct for correlated routes and should always be negative. In case no correlation (and thus no overlap exists), the CF will be zero. The resulting probability of choosing route k from the choice set C is given by:

$$P_k = \frac{e^{V_k + \beta_{CF} CF_k}}{\sum\limits_{l \in C} e^{V_l + \beta_{CF} CF_l}} \tag{4.27}$$

Herein, V_l and V_k are the utility functions of routes l and k, each with a specific commonality factor and the coefficient β_{CF} which is to be estimated. An example of the calculation for the commonality factor is given:

$$CF_k = \ln \sum_{a \in \Gamma_k} \left(\frac{L_a}{L_k} \sum_{l \in C} \delta_{al} \right) \tag{4.28}$$

where L_k and L_a are the lengths of the routes k and a, Γ_k is the set of links belonging to route k, δ_{al} is the link-path incidence dummy which is equal to one if route l uses link a and zero otherwise. The calculation of the available factors however lack a theoretical guidance for which of the different formulations to use.

A different approach is given by Ben-Akiva & Ramming (1998), also presented in Ben-Akiva & Bierlaire (2003) and Ramming (2001): Path Size Logit. Path

Size Logit is based on the theory of Ben-Akiva & Lerman (1985) for aggregate alternatives. The probability of selecting route k is given by:

$$P_k = \frac{e^{V_k + \beta_{PS} \ln PS_k}}{\sum\limits_{l \in C} e^{V_l + \beta_{PS} \ln PS_l}} \tag{4.29}$$

in which PS_k and PS_l are the path sizes of routes k and l and β_{PS} is a coefficient to be estimated. This form thus resembles that of the C-Logit model, but the path size factors are formulated differently.

$$PS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum\limits_{j \in C_n} \delta_{aj}} \tag{4.30}$$

$$PS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum\limits_{j \in C_n} \frac{L^*_{C_n}}{L_j} \delta_{aj}} \tag{4.31}$$

$$PS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum\limits_{j \in C_n} \left(\frac{L_i}{L_j}\right)^{\varphi} \delta_{aj}} \tag{4.32}$$

where Γ_i is the set of all links in path i, L_a ios the length of link a, L_i is the length of path i and $L^*_{C_n}$, δ_{al} is again the link-path incidence dummy and ϕ is a parameter equal or greater to zero.

Ramming (2001) compares the C-Logit and the Path Size Logit (PSL) approaches for estimation results and concludes that the PSL model with the generalized formulation (including φ) outperforms the C-Logit model. Hoogendoorn-Lanser (2005) found that with a value of 14 for φ, the best model fit is found. However, Frejinger & Bierlaire (2007) argue that the original formulation of the PSL model (equation (4.30)) is preferred, as the generalized formulation may produce counter-intuitive results.

The Path Size Logit model is again discussed by Bovy et al. (2008), who derive the correction factor analytically. The Path Size Correction Logit (PSCL) is given by:

$$P_k = \frac{e^{V_k + \beta_{PSC} \ln PSC_k}}{\sum\limits_{l \in C} e^{V_l + \beta_{PSC} \ln PSC_l}} \tag{4.33}$$

in which PSC_k and PSC_l are the path sizes of routes k and l and β_{PSC} is a coefficient to be estimated. Herein, the formulation of PSC_k is given by:

$$PSC_k = \sum_{a \in \Gamma_k} \left(\frac{L_a}{L_k} \ln \frac{1}{\sum\limits_{j \in C} \delta_{al}} \right) \tag{4.34}$$

The PSCL model differs from the PS model with the way the logarithm enters the formulation. The resulting values for the PSC formulation vary between $-\infty$ and 0, whereas the original PS formulation varies between 0 and 1.

Regardless of the specific model, the coefficient for each of the correction factors can be estimated. In the original formulation by Ben-Akiva & Bierlaire (2003) for the PS factor, no coefficient however was included. Ramming (2001) argues that according the theory behind the MNL model, the coefficient should be fixed to one. The estimations for the β_{PS} in his case do significantly differ from both zero and one. Both Hoogendoorn-Lanser (2005) and Frejinger & Bierlaire (2007) argue that the correction factors can have a behavioural interpretation (for example, routes with more overlap can be preferred) and suggest that the coefficients should be estimated. Both also find negative values for estimated coefficients, indeed indicating that alternatives with a higher overlap are preferred over alternatives with less overlap.

With respect to the values for the link lengths it is possible to use the actual length, but also other link attributes related to travel time. Ramming (2001) shows that the PS factor based on the freeflow travel time provides a better model fit. Also Frejinger (2008) applies freeflow travel time for the link attributes. However, neither concludes that the correction factors should be based on lengths or travel times.

Next to the corrections for the MNL model, several models based on the Generalized Extreme Value model (McFadden, 1978) are applied to the route choice situation. Such models are the Paired Combinatorial Logit Chu (1989); Koppelman & Wen (2000) which assumes that routes are chosen among a pair of alternatives within the choice set. This model has not been applied to large size networks. Another model is the Cross Nested Logit model, proposed by Vovsha (1997) and adapted to the route choice situation by Prashker & Bekhor (1998) which assumes routes are chosen within nest which physically correspond to the links in the network. Cross Nested Logit has been applied to full size networks by Ramming (2001), but does not provide satisfying results, neither behaviorally nor computationally. A generalization of the Cross Nested Logit model is presented in Wen & Koppelman (2001); Bekhor & Prashker (2001) and is the Generalized Nested Logit model. This model shows similar results as the Cross Nested Logit models.

Application of Mixed Logit to route choice models concerns both the random coefficients and error components approaches. The first however only accounts for heterogeneity of the drivers regarding the important attributes and heteroscedasticity of the attributes and alternatives. Using a factor analytic specification (see also Walker (2001) for more information on this formulation of Mixed Logit), Bekhor et al. (2002) define a covariance matrix for the error components based on the overlap of all paths in the choice set. They assume the link-specific errors are i.i.d. normal distributed. Frejinger & Bierlaire (2007) adapt this approach by specifying subnetwork components, which does not necessarily mean that two routes using the same subnetwork may share unobserved attributes and have correlation, even without using the same links. For each subnetwork component a covariance parameter is estimated. This model is shown to outperform the PSL model, in terms of model fit. However, as it involves simulation, the time needed for estimation is much higher.

In all of these cases, the route choice models are assumed to be non-adaptive; they are applied to the choice situation at the origin of the trip, and assume that drivers choose a route for this complete trip. However, route choice is made an uncertain environment with many disturbances such as incidents. Traffic information is expected to have a positive effect by enabling drivers to make better route choices as the uncertainty is reduced. However, the previously mentioned models do not reflect the dynamics in the route choice as a driver acquires information on the disturbances for the current trip. This is confirmed by the finding of several authors (Gale et al., 1990; Prato, 2004; Frejinger & Bierlaire, 2007) that drivers choose at anchor points in the road network. Recently, multiple approaches are described which incorporate these sequential choices Gao & Chabini (2006); Gentile & Papola (2006); Gao et al. (2008); Frejinger (2008); Fosgerau et al. (2009). In these approaches, the estimation is not based on the route, but on the individual links. The choice for a route is not interpreted as the choice for a joint route, but as the result of the choice sequence at decision points.

For a route choice model, the individual choices are related to each other, as the choice at a specific junction (apart from the very first junction) is related to the choices at previous junctions. As put forward as an adaptive route choice model in Frejinger (2008), each driver is assumed to choose a route p to the destination at the source node s in the road network, where the node is the source node of the observed link l. This link is part of an ordered set of links L_i, which is observed as i. At each node, there is the individual n and time t specific choice set C_{stn}. In this model, the choice set only contains the three routes each leading to the destination with a minimum travel time, based on the exits at the source node s

The probability of observing the route is then defined as the product of the probabilities of choosing each link l in the observed route, conditional on the arrival of the individual at time t at node s.

$$P_n(i) = \prod_{l \in L_i} P_n(l \mid t, s) = \prod_{l \in L_i} \sum_{p \in C_{stn}} P(l \mid p) P(p \mid C_{stn}; \beta) \tag{4.35}$$

Herein, $P(p \mid C_{stn}; \beta)$ is the route choice model, where β is the vector of coefficients to be estimated in this model. $P(l \mid p)$ is a dummy, which equals one if link l is the first link in the route p, and zero otherwise. This is a so-called adaptive route choice model, as it includes the possibility for a traveller to be adaptive in his/her route choice, by re-evaluating the routes at each intermediate decision point. As pointed out by Gao & Chabini (2006); Frejinger (2008); Gao et al. (2008), this does not take into account the future states of the road network, as these are expected to be time dependent too. They present an approach taking the future possible states into account, called a routing policy.

These approaches however do need the alternatives at each decision point, for which a choice set needs to be generated. As the choice set is time dependent, the generation is more complex and expected to have bigger impact on the choice itself. Especially in case of provision of information the choice set assumptions are of importance, as the driver can be aware of the actual travel times, in stead of knowing a historic (perceived) based travel time.

4.3.4 Proposed exit choice model

For this research we assume a driver makes a choice at each decision point in the road network, choosing one of the direct alternatives at that location. These are the links which have the current decision point as the source node, excluding the link which concerns a u-turn. In general, no more than three alternatives are available at each decision point. We label these alternatives according to the angle with the entry link (the link having the decision point as the sink node); right, straight ahead or left. We further assume that the choice at each decision point is independent of the previous decision point; in other words, the route towards the current decision point does not affect the current choice for the exit or following choices. As a driver does need to get to his destination, we further assume that the shortest path (based on actual travel times) is used for the route starting with the exit at the current decision point. The choice set at each decision point thus consists of a maximum of three labelled alternatives and is time dependent. The alternatives will always have a certain amount of overlap, which will always occurs at the end of each route. In order to account of this overlap it is possible to include a correction factor, for which we propose to use the PS or PSC approaches.

The decision rules leading to the choice model can be based on several choice models mentioned above, amongst which the MNL and mixed logit models. Which of these models performs best and provides for consistent estimates will be answered in chapter 9.

4.3.5 Route choice factors

Next to the models which are needed to estimate the choice for a route, the attributes of the routes or alternatives are important. An overview of attributes which are of importance is given in table 4.1 based on the report of Van Dijck (2007). Each of these main characteristics is discussed briefly.

Driver characteristics

The characteristics of the driver are an important aspect for all time instants where a route choice is evaluated. For example, age and income show a significant impact on the route choice for a morning commute (Li et al., 2005). Srinivasan & Mahmassani (2003) focused on the heterogeneity for the propensity of switching routes under information provision, based on effects of age and gender. Habit or inertia is often investigated and expected to be an important aspect in route choice, as drivers are less likely to optimize their route choice in case of a habitual route (see for example Polydoropoulou et al. (1996); Srinivasan & Mahmassani (2000); Gärling & Axhausen (2003); Peeta & Yu (2005)). Bogers et al. (2005) also include the learning behaviour of a driver, which also relates to the experiences with traffic information en route and post trip. Learning also has received much attention in the field of route choice (Adler, 2001; Klügl & Bazzan, 2002; Ben-Elia & Shiftan, 2010). The risk attitude of a driver is related to the time constraints, or the utility

Table 4.1: *Factors influencing route choice behaviour (Van Dijck, 2007)*

	Pre trip	**En route**	**Post trip**
Driver characteristics	• Socio economic • Trip purpose • Inertia & Habit • Compliance • Time constraints • Risk attitude • Learning behaviour	• Experience with traffic information • Experience with network • Location in the road network	• Experience with traffic information • Experience with network
Network characteristics	• Feasible paths • Non-recurrent conditions • State of network • Design of network	• Feasible paths • Non-recurrent conditions • State of network	
Traffic information characteristics	• Source	• Source • Information strategy • Quality • Cost	

related to early or late arrival, due to variability in the travel time (Noland & Polak, 2002; Bonsall, 2004b).

Network characteristics

The lay-out or design of the road network obviously influences the route choice and the amount of alternatives. If just two routes between origin and destination are available, a driver has not much to choose from. In an urban road network, the amount of routes is much larger, but not all routes are as likely to be used. The design of the road network, with respect to the junctions, traffic lights, highways, one-way roads, speeds, etc. can greatly influence the route choice. For example, the amount of left-turns or speed bumps were significant in the route choice models of Frejinger (2008). Also the current traffic state is of importance, for example with respect to the travel time variability, or incidents. In general, the non-recurrent conditions that occur are expected to influence the route choice of a driver.

Traffic information characteristics

With respect to the traffic information characteristics, the source of the traffic information is of influence on the route choice. For example, a Variable Message Sign (VMS) provides different information than a personal navigation device, or traffic information received via the radio. A VMS will only provide information on a certain location in the road network, whereas a navigation system continuously provides a route advice (based on current or freeflow traffic conditions). The characteristics of the information, such as quality, cost and type of information are also very important (Van Berkum & van der Mede, 1993; Adler & Blue, 1998; Chen et al., 1999; Srinivasan & Mahmassani, 2003; Jou et al., 2005; Chorus et al., 2006b).

4.4 Summary

Discrete choice modelling is a popular modelling approach in the context of traffic and transport. Discrete choice modelling is based on assumptions regarding the decision maker, the alternatives, the attributes hereof and a decision rule. One of the most used assumptions is that of random utility maximization, which assumes a decision maker will acquire a certain (dis)benefit of choosing an alternative and which he or she tries to maximize (or minimize). Since the analyst cannot observe all the characteristics of the alternatives, a random error is introduced in the utility function. The most common (and most restrictive) approach for modelling the choice probability is the Multinomial Logit (MNL) model, which assumes all random terms to be independent and identically Extreme Value Type 1 distributed which leads to the following probability of individual n to choose alternative i from the set of alternatives C_n:

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum\limits_{j \in C_n} \left(e^{\mu \mu V_{jn}}\right)} \tag{4.36}$$

Estimating the systematic utility functions V_{in} can be done using maximum likelihood.

The MNL model has advantageous properties, as it is easily estimated using available software packages and is flexible in the formulation of the utility functions. The results of the estimation are often used and provide consistent estimates and are easily explained. However, the assumption of independent and identically distributed random terms lead to the property of Independence of Irrelevant Alternatives. This property means that the ratio of the probability of any two alternatives is unaffected by any other alternatives, which not holds true if there is correlation in the alternatives. The famous "Red Bus-Blue Bus" is a clear example hereof. There are several approaches to relax the underlying assumption. We presented the Nested Logit, Heteroscedastic Extreme Value Logit and Mixed Logit approach.

Route choice is presented as a distinctive case of discrete choice modelling, as the choice set is fairly large which is not expected to be complete considered by the decision maker. Next, the alternatives in the choice set are likely to be correlated (have overlap). To this end, several methods have been presented that generate an appropriate choice set, and various models that are dealing with the correlation of these alternatives.

Based on these models, we propose a route choice model which is based on distinct alternatives at a decision point: the exit links. We assume a driver makes a decision at each decision point, contrary to a choice at the start of the trip. This enables us to model the route choice during the trip as well. The alternatives are described using the exit link as label (based on the angle with the entry link) and the current fastest route to the destination starting at this exit link. As the alternatives are based on the current situation and decision point, the choice set is time and location dependent. With respect to the decision rules, we assume the familiar MNL model, together with a correction factor for the possible overlap in the alternatives.

In chapter 5 a stated preference research using random utility maximization is used, to investigate the choice of drivers for different types of information, together with three important attributes of this information. The proposed exit choice model will be estimated using data from a route choice simulator, also using the random utility maximization approach. The estimated models are presented in chapter 9.

Some prefer carrot while others like cabbage.

— Irish saying

Chapter 5

User preferences for traffic information

5.1 Introduction

This chapter presents the approach and results of a survey undertaken to study the preferences of drivers for traffic information. The results of the survey provide an answer to research question 2: *What traffic information do drivers prefer in non-recurrent situations?* This research question has several aspects related to it. We are interested in differences caused by varying traffic situations, but also by varying types of drivers. We also study the choice of varying attributes of traffic information and the differences between types of drivers.

We start with providing a conceptual model which relates the attributes and contents of traffic information to the situations and individual characteristics (section 5.2). Based on the conceptual model we designed a internet survey, including a stated preference study related to discrete choice modelling. The results of the survey are presented in section 5.3. We finish the chapter with a discussion (section 5.4) and a summary (section 5.5).

5.2 Approach

5.2.1 Conceptual model

The conceptual model describes the relationships between the following elements that influence the preferences for traffic information as presented in figure 5.1:

- situations;
- personal characteristics;
- information attributes;
- information types;
- information contents.

Three elements are of specific interest, in terms of preferences of drivers. These are the information attributes, information types and information contents. When looking at the conceptual model provided, these preferences are influenced by the

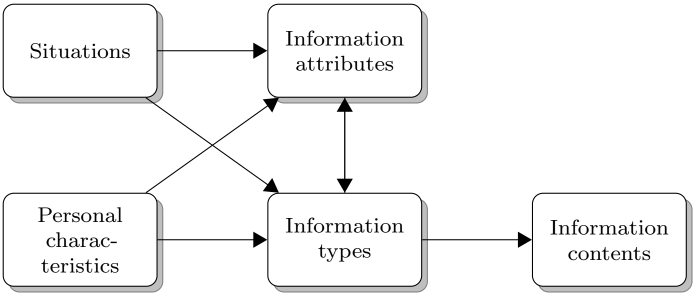


Figure 5.1: *Conceptual model*

two other factors. However, to explain the internal relationships that are also involved it is necessary to further describe the five elements that are used.

Situations describe the context of travel, for example the road network, time of day, motive and specific traffic situations. The three important non-recurrent situations (accidents, road works and large events) are contained within this element. Personal characteristics describe the traveller, which include, for example the standard socio-economic or socio-demographic values as well as familiarity with traffic information, attitude towards mobility and current usage of traffic information.

The conceptual model distinguishes three related elements concerning traffic information:

- attributes;
- types;
- contents.

Attributes of traffic information concern things as costs, accuracy and reliability. Types can be defined as descriptive or prescriptive, whereas contents deal with the actual information provided to the traveller, such as length and location of congestion, or advice on the fastest route. For all three elements, a traveller is expected to have a certain preference, based on the situation and personal characteristics. For both contents and types, the attributes are important too. Travellers are expected to have preferences for levels of these attributes, and these attributes influence the preference for certain types of traffic information.

The situations and the personal characteristics are expected to influence the preference of travellers (or drivers) for types of traffic information and their preference for (levels of) attributes of traffic information. These attributes again influence the preference for types of traffic information, but this relationship also works the other way around; the type of traffic information influences the preference for attributes. The preference of drivers for types of traffic information determine the preference for certain contents of traffic information.

Based on the conceptual model, several hypotheses can be made. For these hypotheses, specific elements have also been selected for the personal characteristics and situations. To start with, it is expected that for the information contents (or information types) a number of socio-economic and socio-demographic variables influence the preferences for traffic information. Especially age, gender, level of education, current usage of traffic information and attitude towards more mobility are expected to have significant influence. In this case, attitude towards mobility can be a proxy for other socio-economic or socio-demographic variables, which is partly suggested by Need (2002). It is also expected that the situation influences the preferences for contents or types traffic information in a significant way. Of course, the traffic situation encountered is part of the situation and expected to show a very important relationship. Trip purpose is also expected to be significant with respect to the preference for type and contents of traffic information.

Concerning the preference for attributes of information, it is expected to be related to the types of information, the personal characteristics and the situations. The same basic elements for all relations will be used here, as for the investigation of preferences of types of traffic information.

5.2.2 Survey design

Based on the findings in chapter 2, as well as the conceptual model, a survey was expected to be the best available method to acquire results to test the mentioned hypotheses. This section will describes the development and design of this survey.

In total the survey consisted of five sections:

1. general information - background information of the respondent, such as age, car possession, most used motive for car use, etc.
2. attitude towards mobility
3. current traffic information usage - is the currently available traffic information used, which sources for traffic information are used, and to what extent are users satisfied with this information?
4. traffic information contents - which contents of traffic information are chosen in different situations?
5. traffic information attributes and types - which attributes for different types of traffic information are important?

General information

Several socio-demographic and socio-economic variables are expected to be important as determinants for the preferences for traffic information. The survey therefore started with questions on these topics for all respondents. This also provided a selection question to allow only respondents with driving licenses. There were several questions on age, gender, level of education, household size and possession of vehicles. Respondents were also asked about the amount of kilometers they drive per year and the distribution of these kilometers over highways, regional roads and urban roads.

Attitude towards mobility

This section of the survey was dedicated to the respondents' attitude towards mobility, based on the definitions of Need (2002). It contains 17 statements on which the respondent has to indicate to which extent he agrees. A five point scale is used for this purpose, together with the "no opinion" and "not applicable" options. A given rule set determines which respondent belongs to which group.

Current traffic information usage

It was expected that familiarity with traffic information could be a major issue in preferences for new or other types of traffic information. Therefore this section was added to the survey. It also enabled an overview of the most used sources for traffic information. Respondents were asked to indicate if they possessed a navigation system (and which type) and whether they use traffic information before and during travelling. If they did, they were also asked to provide the source for traffic information they used most often. The section ended with the satisfaction of respondents on the traffic information they used, as this could be an indication for improvements in traffic information.

Traffic information contents

This section of the survey concerned the last element in the conceptual model (see figure 5.1). As given in this model, the preference for contents of traffic information is expected to vary over travellers and situations. The personal characteristics of a traveller have been mentioned in the three previous sections and are expected to be the same in different situations. A situation however, can change and thus more directly relates to the preference for contents of traffic information. Because of this direct expected relation, situations were used as a starting point for determining the preferences for traffic information contents.

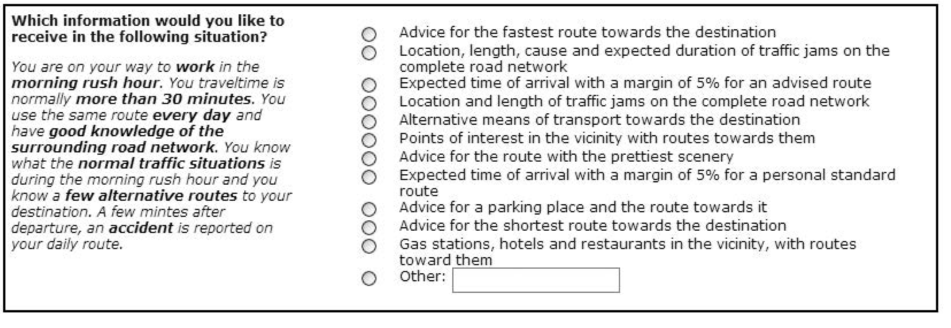
The starting point for the situations are the three non-recurrent traffic situations: (a) accidents, (b) road works and (c) large events. Other elements that described the situation are, for example, trip purpose or motive (also changing, compared to personal characteristics), time of day, and familiarity with the network and traffic situations that can occur. Because of certain relationships within the situations (for example, a large event relates to a recreational motive and road works cause unfamiliarity with the traffic situations) not all possible combinations were used. The situations were described textually to the respondent. They were asked to choose as if they were in such a situation. An example of such a question is given in figure 5.2.

The respondents could choose from eleven pre-defined contents for traffic information. These eleven options were created following an analysis of currently available traffic information from radio bulletins and navigation systems, together with discussion on this subject with colleagues. The eleven options were:

- location and length of traffic jams on the complete road network;
- location, length, cause and expected duration of traffic jams on the complete network;
- advice for the fastest route towards the destination;
- advice for the shortest route towards the destination;
- advice for the route with the prettiest scenery;
- advice for a parking place and the route towards it;
- expected time of arrival with a margin of 5% for a personal standard route;
- expected time of arrival with a margin of 5% for an advised route;
- gas stations, hotels and restaurants in the vicinity, with routes towards them;
- points of interest in the vicinity with routes towards them;
- alternative means of transport towards the destination.

The options did not provide the actual information, but were a short description of it. To minimize errors caused by respondents always choosing the same option, the order of the eleven options was randomized between questions. Respondents were forced to make deliberate choices regarding the contents of traffic information.

In this part of the survey, the attributes of the traffic information were not considered. This also includes the presentation device and format. All contents can be used on various devices such as navigation systems and mobile phones. The presentation format, although important, is not included in the survey. This can change from graphical, textual or audio.


Which information would you like to receive in the following situation?

You are on your way to ***work*** *in the* ***morning rush hour****. You traveltime is normally* ***more than 30 minutes****. You use the same route* ***every day*** *and have* ***good knowledge of the surrounding road network****. You know what the* ***normal traffic situations*** *is during the morning rush hour and you know a* ***few alternative routes*** *to your destination. A few mintes after departure, an* ***accident*** *is reported on your daily route.*

- Advice for the fastest route towards the destination
- Location, length, cause and expected duration of traffic jams on the complete road network
- Expected time of arrival with a margin of 5% for an advised route
- Location and length of traffic jams on the complete road network
- Alternative means of transport towards the destination
- Points of interest in the vicinity with routes towards them
- Advice for the route with the prettiest scenery
- Expected time of arrival with a margin of 5% for a personal standard route
- Advice for a parking place and the route towards it
- Advice for the shortest route towards the destination
- Gas stations, hotels and restaurants in the vicinity, with routes toward them
- Other:

Figure 5.2: *Example of the survey for information contents*

Traffic information attributes and types

The last part of the survey considered the relationship between information types and their attributes. Situations should be incorporated in this section, but because of practical reasons, just one "normal" situation has been used to prevent respondent fatigue. This part of the survey is based on the theory of discrete choice modelling, presented in chapter 4.

For the design and modelling of the choice for different types of traffic information and different attributes hereof we assume the following utility function:

$$\begin{aligned} V_i = & ASC_i + \beta_{cost} Cost_i + \beta_{timeliness} Timeliness_i \\ & + \beta_{reliability} Reliability_i \end{aligned} \tag{5.1}$$

where:

ASC_i =Alternative specific constant of alternative i

β_k =Coefficient for attribute k

Equation (5.1) provides the attributes that are used in this study, namely:

- costs;
- timeliness, and;
- reliability.

These attributes were chosen based on the findings in chapter 3. No other attributes were chosen because it would greatly increase the size of the stated choice part of the survey. Costs concern the purchase of a device capable of receiving traffic information and if necessary providing advice. It also includes costs of receiving the traffic data, when this information is transmitted using the GSM network. Timeliness concerns the delay between the measurement of data at the road and the moment the traffic information is received. In this period, all kinds of calculations and transformations of the data can take place, in order to create usable information. Reliability shows the level to which the traffic information agrees with the actual situation the driver experiences on the road. For each attribute three levels were defined. These can be found in table 5.1. The number of three levels per attribute is of influence on the design of the stated choice experiment.

Besides the attributes and their levels it is also necessary to define the alternatives which could be chosen by the respondents in the single situation. It would be advantageous for analysis to also use the same eleven alternatives as presented earlier (see item 5) because in both cases, the alternatives are the same. However, this would make a very difficult choice for respondents, because they have to evaluate 11 alternatives based on their contents and their attributes before they can choose. Therefore the alternatives have been condensed into four types of traffic information and the opt-out option (chosen because of the introduced costs for alternatives). This results in the following five types:

Informing information information on the length and location of congestion, and possibly causes

Table 5.1: *Attribute levels and their descriptions*

Attribute	Level	Description
Costs	-1	€ 0.10
	0	€ 0.55
	1	€ 1.00
Timeliness	-1	No delay
	0	Around 5 minutes delay
	1	Around 15 minutes delay
Reliability	-1	Is sometimes correct
	0	Is often correct
	1	Is always correct

Advising information information in the form of a general advice

Personalized information information in the form of a personalized advice, based on the important personal criteria

Predictive information information based on the personalized information alternative, but also including predictions of future traffic situations

No information no information at all

Only the first four alternatives have attributes, because it is not possible to provide costs, timeliness or reliability to the "no information" alternative.

The next step in the development of the stated choice experiment is the design of the experiment. In other words, which combinations of attribute levels are used to develop the alternatives used in questions, together with the number of questions used in the experiment. Several methods are available to design this experiment, but based on the approaches presented in Louviere et al. (2000) and Hensher et al. (2005), a fractional factorial design was chosen. In a full factorial design, all the available combinations of attributes and attribute levels are used to develop the choice questions. The more alternatives, attributes and attribute levels are used, the larger the design will be. For this study, a full factorial design would lead to having 531,441 questions (based on three levels, and 12 attributes, which leads to 3^{12} combinations). The full factorial design allows to estimate all the interaction effects that might exist (simultaneous effects of two attributes, for example an interaction between costs and reliability). However, because the number of questions is very large, it burdens the respondents. In order to reduce the number of questions, a fractional factorial design is used.

For a fractional factorial design it is necessary to select the interaction effects that are of interest, based on the utility function that needs to be estimated. In this thesis no interaction effects are used, which resulted in 25 degrees of freedom (based on $(L-1)MA+1$ where L is the number of attribute levels (3), M the number of alternatives (4) and A the number of attributes (3)). In a fractional factorial design, this results in an orthogonal array of 27 rows by 13 columns. The rows are treatment combinations (or questions) of varying levels of attributes for all alternatives. The columns represent the attributes for the alternatives and one column which is used for blocking. Blocking means that the number of treatment combinations presented to a respondent can be reduced. Three blocks of 9 questions are used to question the respondents. In order to maintain orthogonality in the analysis of the results, it is necessary to have an equal number of blocks with data.

5.2.3 Survey method

The survey was designed to be an internet survey for various reasons. First it allowed for special routing and randomization of the questions, which was necessary for the part in the characteristics of traffic information. Second, an internet survey allows for easy collection of the data because it is already stored in a database. The data does not have to be manually entered into a database. Third, it provides a quick and cheap method to collect the necessary data.

Based on this, a choice was made to use so called Open Source software. A large number of systems are available as Open Source, but based on functionality and adaptability, as well as good documentation, Mod_Survey was selected. Mod_Survey is licensed under the GPL v2 (GNU General Public License, see http://www.gnu.org/copyleft/gpl.html) and can be found at http://www.modsurvey.org/. Mod_Survey is a modperl module for Apache and can be run on any webserver that runs these two programs. The final survey was set up on http://www.kenniscentrumaida.nl/enquete/survey.survey. The resulting survey as it could be viewed online can be found in appendix A.

The estimation of the various discrete choice models will be done by using Biogeme (Bierlaire, 2003, 2008).

The survey was targeted at the Dutch owner of a driving license. Respondents were able to take part if they did not possess a car, as long as they possessed a driving license. The survey did not aim at the professional (truck) driver, as this is beyond the scope of the total research project. Professional drivers could take part, but from a personal point of view.

The sample strategy used for this study is Simple Random Samples (Louviere et al., 2000). In a simple random sample study, each member of the population has an equal chance of being selected for the sample. For each of the selected individuals, the personal characteristics and choices are recorded. This means that the sample will be more or less an equal representation of all the groups in the population.

Sample strategy

Based on the sample strategy mentioned above a sample size can be calculated. Two different formulas were used for this calculation. The first is more generic, while the later is specifically based on the part of the information attributes (the stated choice part). The first equation is:

$$n = z^2 \frac{p(1-p)}{a^2} \tag{5.2}$$

in which:

$n =$ sample size
$z =$ standard score for the given confidence level a
$p =$ an estimate of the proportion falling into the group of interest
$a =$ accepted error

This equation (Buijs, 2008) can be used when the population N is large compared to the sample n because of the use of the normal distribution. In this case equation (5.2) may be used because the population is almost ten million. When using a confidence level of 95%, z becomes 1.96. The most conservative value for p is 0.5 which would result in the largest sample size necessary. The accepted error was set on 0.05. The resulting sample size is 385.

The second equation is:

$$n = \frac{1-p}{rpa^2}\left[\Phi^{-1}\left(\frac{1+\alpha}{2}\right)\right]^2 \tag{5.3}$$

in which:

$\Phi =$ inverse cumulative normal distribution function
$\alpha =$ accepted error level
$a =$ allowable deviation between true and estimated choice proportion
$r =$ the number of questions per participant

Equation (5.3) is applied to the stated choice part of the survey. Equal proportions for all 5 options have been assumed, which means $p = 0.2$. α is set on 5%, while the accepted error a is the same as in the previous equation. The number of questions per respondent is 9. This results in a sample size of 683 respondents. Results for different accepted error levels are presented in table 5.2. With both values in mind, a sample of at least 700 respondents seems to be appropriate for at least the stated choice part.

Table 5.2: *Sample size with different accepted error levels*

Error level	Sample size
5%	683
6%	475
7%	349
8%	267
9%	211
10%	171

5.2.4 Invitation, response and check

Respondents for the survey were found in two ways. One was via a market agency with an internet panel. The other way used personal and business contacts. The market agency that was used was RM Interactive, which has an internet panel of more than 25.000 members. RM Interactive has to date performed more than 1000 online surveys.

RM Interactive agreed to guarantee at least 300 completed surveys, for which a contract was signed. The members of the internet panel received 2 euros for filling in the survey. Apart from that, they could win a prize (20 euros) with a 1/25 chance. RM Interactive selected members of their panel based on the target group of "the Dutch car driver". The members of the panel were invited by email to participate in this research with a personal link to the survey. The invitation for members of the panel was open from 31 October to 22 November (2005). Halfway through this period a reminder was sent to the invited members who had not yet completed the survey.

The second way of finding participants consisted of sending an email with a link to the online survey to various groups, such as family, friends, colleagues and business contacts. They were all asked to forward the invitation to participate in the survey to their own contacts (a snowball method). The invitation to participate in the survey was also placed on the intranet of research organization TNO and the website of Knowledge centre AIDA. The survey was online from 21 October 2005 to 8 January 2006.

Before the data were analysed, a check was performed on the integrity of the data. First the origin of the data, in the form of an IP address was examined. It seemed as if some respondents had sent the data twice or even more times. The answers of these respondents were checked. It showed that they either came from a company where a proxy was used, or a private address where a family had filled in the survey, giving more than one respondent per IP address. No data had to be deleted.

The next check on the data concerned the time respondents needed to fill in the survey. The average time needed was more than 20 minutes, which is due to the large amount of time that was registered for some respondents. This is also indicated by the median, which was around 14 minutes. The complete data set was divided in two groups. One of the groups consisted of answers with a fill out time of less than 600 seconds (which consisted of 141 sets of answers), while the other group had a fill out time of more than 600 seconds (454 answer sets). The answers of these two groups were compared with respect to the basic sample characteristics such as gender, age and education. It showed that the characteristics from the group with a low fill out time did not differ significantly on most characteristics from the other group (the characteristic that differs significantly is household size). Because of this and the need for as much data as possible, all the answers were used in further analysis. This resulted in a total data set of 598 answers, apart from a few questions where a respondent did not provide any answers. These failing answers were set to missing values in the analysis.

5.3 Results

5.3.1 Participants

In total 598 respondents filled in the survey. The sample was compared to the population of Dutch owners of a driving license, (taken from Statline (CBS, 2011)), in terms of age, gender and level of education (see table 5.3). The respondents had an average age of 42, and most respondents had a higher level of education. 372 of the respondents were male and 226 were female. Based on the results of Chi-Square tests for age, gender and level of education, the sample differed significantly from the target group ($p<0.001$).

More background information on the sample is provided in tables 5.4 and 5.5. Most respondents live in a household of three or more persons. The average respondent has had a driving license for 21 years and his or her own car. 15.1% of the respondents possessed a lease car, which is high compared to the Dutch average of 7%. Most respondents drive between 10,000 and 20,000 kilometers each year, which is in line with the Dutch average of 15,500 kilometers per year. Owners of a lease car, however, drive many more kilometers per year. This also coincides with the respondents having a business motive for car travel. However, most respondents used their cars for private travel.

5.3.2 Attitude towards mobility

Part of the description of the sample is the attitude towards mobility (see section 3.4 for a presentation of the different attitudes). Table 5.6 shows a comparison of the sample and the population in percentages. The population is taken from Need (2002); Ministerie van Verkeer en Waterstaat (2004) where 1832 respondents took part in the survey. The results in their study were then weighed by age, gender, level of education, social class, household size, province and degree of urbanity, to acquire a proper representation of the total Dutch population. According to the

Table 5.3: *Characteristics of sample and population*

Characteristic		**Sample**	**Population**
Gender	male	62.2%	54.2%
	female	37.8%	45.8%
Age	18-24	6.7%	8%
	25-44	51.3%	44%
	45-64	37.3%	36%
	65 and older	4.7%	12%
Level of education	primary education	1.0%	4%
	lower secondary education	7.7%	24%
	higher secondary education	33.6%	44%
	higher education/university	57.0%	28%
	other/unknown	0.7%	-

Table 5.4: *Characteristics of sample (1)*

Characteristic		#	%
Household size	one person	104	17.4%
	two persons	224	37.5%
	three persons or more	270	45.2%
Car possession	none	57	9.5%
	one or more	541	90.5%
Lease car possession	none	457	84.9%
	one or more	81	15.1%
Total kilometers	0-10,000	160	26.8%
	10,000-20,000	225	37.6%
	20,000-30,000	114	19.1%
	more than 30,000	99	16.6%
Motive for car use	commute	217	3.,3%
	business	96	16.1%
	private	285	47.7%

Table 5.5: *Characteristics of sample (2)*

Characteristic		Mean	SD
Years of possessing driving license		21 years	13 years
Percentage kilometers per road type	highway	45	24
	provincial road	28	18
	urban road	25	18

Chi-Square test, the sample differs significantly from the population (p<0.001). The sample has a large group of respondents that can be described as Deliberates.

The attitude towards mobility for the sample has been investigated for the basic characteristics of the sample. The resulting table can be found in table 5.7. In the different groups, no significant difference in gender can be found according to the Chi-Square test. However, for age and level of education, significant differences exist in the sample. For age, the group of 25 to 44 is over-represented for the group of Deliberates, while for the ages of 45 to 64 and 65 and older, the Conscious and Enjoyers are over-represented. For the level of education, the group of higher or university educated respondents largely makes up the group of respondents with the attitude Deliberates. In general, one can assume that the over-representation of the Deliberates in the sample is caused by a high number of young higher educated respondents.

Table 5.6: *Attitude towards mobility*

Group	Sample %	Population %
Accepters	29.9%	26%
Deliberates	28.1%	18%
Conscious	13.3%	17%
Competitors	15.2%	21%
Enjoyers	13.4%	18%
No group	6.7%	11%

Table 5.7: *Attitude towards mobility for basic characteristics*

Characteristic		Accepters	Deliberates	Conscious	Competitors	Enjoyers	No group
Gender	male	64%	59%	69%	61%	61%	58%
	female	37%	41%	31%	39%	39%	43%
Age	18-24	5%	11%	7%	7%	3%	3%
	25-44	53%	59%	52%	32%	45%	58%
	45-64	38%	29%	35%	47%	45%	38%
	65 and older	3%	1%	6%	14%	7%	3%
Level of education	primary education	1%	1%	0%	1%	3%	0%
	lower secondary education	5%	5%	7%	11%	12%	15%
	higher secondary education	35%	23%	35%	41%	43%	38%
	higher education/university	58%	71%	55%	47%	43%	48%
	other/unknown	1%	1%	2%	0%	0%	0%

5.3.3 Current information usage

According to available research, familiarity with traffic information plays in important role in the understanding and acceptance of new types of traffic information. It can also influence the preferences for traffic information. This section presents the results of the survey, regarding the familiarity with the currently available types of traffic information in the Netherlands.

In tables 5.8 and 5.9 the results of questions on this topic can be found. Already 26% of the sample possess a navigation system (in 2006), which would mean that around 2 million vehicles in the Netherlands will be equipped with such a system (if the group of respondents is representative for the Dutch population). However, most of the systems are portable which means that these can be used in more than one vehicle. Quite a lot of the navigation systems are, however, not portable.

Most respondents use information both before and during their trip, although during the trip more information is used. Both before and during their trip, the radio is the most used source of traffic information. During the trip, the radio is even more important as source for traffic information, probably because all vehicles are equipped with a radio. Before travelling, teletext and internet are also important sources of traffic information. Both provide “complete” information, by presenting a list of all known traffic jams, or even showing a map of all the locations where traffic jams occur (internet). Often mentioned by the respondents is the website of the Royal Dutch Touring Club, the ANWB, which provides such a graphical overview of congestion.

During their trip, respondents also show a high usage of RDS-TMC. It is expected that this high percentage is caused by a wrong interpretation of the term RDS-TMC by the respondents. In all modern car-radios a system is available which switches channels when a traffic information bulletin is available (this system is called RDS-TA, which stands for Traffic Announcement). RDS-TMC however, does not switch channels, but uses available bandwith to send text messages with

Table 5.8: *Current usage of traffic information*

Question	Answer	#	%
Possession of a navigation system	yes	155	25.9%
	no	443	74.1%
Information usage before travelling	yes	356	58.5%
	no	242	40.5%
Information usage during travelling	yes	446	74.6%
	no	152	25.4%
Satisfaction over used information	very satisfied	22	3.9%
	satisfied	312	55.7%
	neither satisfied or unsatisfied	111	19.8%
	unsatisfied	68	12.1%
	very unsatisfied	5	0.9%
	no opinion	42	7.5%

Table 5.9: *Available and used information sources (multiple answers possible)*

Question	Source	#	%
Type of navigation system	PDA	47	7.9%
	Smartphone	6	1.0%
	TOMTOM (Go)	28	4.7%
	In-car (from production)	46	7.7%
	In-car (retrofitted)	27	4.5%
Information source pre trip	Radio	259	43.3%
	Teletext	166	27.8%
	GSM or SMS	10	1.7%
	Navigation system	26	4.3%
	RDS-TMC	69	11.5%
	Internet	110	18.4%
Information source during the trip	Radio	410	68.8%
	Teletext	3	0.5%
	GSM or SMS	15	2.5%
	Navigation system	51	8.5%
	RDS-TMC	113	18.9%
	DRIP	84	14.0%
	Internet	5	0.8%

the length and location of congestion. RDS-TMC is mainly used in navigation systems. A third rather important source of information are the DRIP's (Dynamic Route Information Panel). These are large text displays above the highway at certain locations where two or more routes to one destination are available. The panels show the amount of congestion or the travel time to the destination.

It is also interesting to see the satisfaction of the used information. Most respondents indicated that they were satisfied or very satisfied with the traffic information they used. This contradicts the often heard complaints of traffic information, especially concerning radio bulletins, that the traffic jam they are in is not mentioned on the radio bulletins. The assumption that travellers are looking for new and other types of traffic information because of a low satisfaction over the currently available information is therefore not valid.

The current usage of traffic information can also be looked upon in more detail by looking at the relationship between the usage of traffic information and the other sample characteristics. By using contingency tables (see appendix B) and Chi-Square statistics it is possible to identify the sample characteristics that significantly influence the possession of a navigation system, and the usage of traffic information before and during travel. The sample characteristics taken into account are: (a) gender, (b) age, (c) level of education, (d) kilometers per year, (e) motive for car use and (f) attitude towards mobility .

For gender, a significant influence is found for the usage of traffic information, both before and during travel. Males use more traffic information before and during their travel, which might be due to the fact that more males work away from home compared to females. For age, also a significant difference has been found between the age groups compared to the usage of traffic information before and during travel. No difference was found for the possession of a navigation system. The older a respondent is, the more likely he or she is to use traffic information, both before and during travelling. This might be due to the fact that the younger respondents are mostly students, who do not travel that much by car, and thus have a lower need for traffic information. For the level of education only a significant difference was found for the usage of traffic information during travel. Respondents with a higher education or university background are more likely to use traffic information during their trip, compared to the other levels of education.

The amount of kilometers a respondent drives per year showed a significant relationship to the possession of a navigation system. The more kilometers one drives, the more navigation systems are owned (and used). The same relationship was found for the usage of traffic information during travel, but not before travelling. The same relationships were found for the most mentioned motive for using a car. The respondents with a business motive, have a higher degree of possessing a navigation system. Before the trip, the motive does not significantly influence the usage of traffic information. During the trip, again the business motive shows a higher usage of traffic information, compared to the commute and recreational motive. Last but not least, the attitude towards mobility shows a relationship to the usage of traffic information before travel. Enjoyers as well as Accepters show less usage of traffic information before their trip, while the Conscious show a higher usage. For Enjoyers and Accepters it is easy to see that because they are not really interested in optimizing their travel, there is less need for traffic information. Conscious, on the other hand want to optimize their trip in terms of safety and environment, and thus seek detailed information on the trip to be made.

5.3.4 Traffic information contents

Preferences for contents of traffic information were expected to be related to both situational and personal characteristics (see section 5.2). This section will present the results of the survey for which eleven situations were defined. The situations were aggregated based on specific characteristics that define the situations. This means that there still are differences with other characteristics, which could influence the results. Three characteristics will be presented here, which are (a) non-recurrent traffic situations, (b) motives and (c) familiarity. But first the total results for all eleven situations combined will be discussed (these can be found in tables 5.10 to 5.12). Most respondents preferred to have traffic information in the form of advice on the fastest route, irrespective of the traffic situation at hand. Second and third most chosen were extended traffic information (location, length, cause and expected duration of traffic jams) and expected time of arrival with a margin of 5% for an advised route. Striking is the large amount of respondents who chose advice for the shortest route. This can be explained when respondents do not see a difference between the shortest and the fastest route. The most preferred traffic information is thus related to travel time.

Non-recurrent traffic situations and preferences

In section 5.2 we hypothesized that the traffic situations that (can) occur during a trip influence the preferred contents for traffic information. Therefore, the survey presented several situations to the respondents, which varied according to the specific type of non-recurrent traffic situation that occurred during that trip. Respondents had to choose among eleven options for their preferred contents of traffic information. The total results are displayed in table 5.10, for all eleven situations combined and for the three combined non-recurrent situations. According to the Chi-Square test the three non-recurrent traffic situations are significantly different from each other and all the situations combined ($p<0.001$).

For traffic situations with accidents, respondents have a greater preference for extended traffic information, compared to other situations. Drivers do like to know the cause and expected duration of traffic jams due to the accident, and want to be able to make their own route choice based on this information. During road works, however, respondents show a greater preference for advice on the fastest route and a much lower preference for extended traffic information. For the situation with a large event, the respondents clearly have a higher preference for other types of information, for example advice for a parking place or advice for alternative means of transport. For all three situations, the specific characteristics of the non-recurrent event influence the preference of drivers. For accidents, deviation from the average situation seems to be important. For large events, people often

Table 5.10: *Traffic information contents for different traffic situations in % and absolute numbers*

	All situations	Accident situations	Road works situations	Large event situations
Location and length of traffic jams on the complete road network	6.6% 421	7.8% 182	5.4% 158	6.9% 81
Location, length, cause and expected duration of traffic jams on the complete network	19.7% 1261	24.6% 572	16.1% 468	18.8% 221
Advice on the fastest route towards the destination	33.3% 2134	29.6% 689	36.4% 1056	33.1% 389
Advice on the shortest route towards the destination	10.9% 699	10.4% 242	12.6% 367	7.7% 90
Advice on the route with the best scenery	3.2% 204	3.2% 74	3.4% 98	2.7% 32
Advice on a parking place and the route towards it	1.3% 86	0.6% 14	0.5% 14	4.9% 58
Expected time of arrival with a margin of 5% for a personal standard route	6.7% 432	6.0% 140	7.3% 213	6.7% 79
Expected time of arrival with a margin of 5% for an advised route	13.6% 871	14.3% 332	13.8% 400	11.8% 139
Gas stations, hotels and restaurants in the vicinity with routes towards them	0.8% 49	1.1% 26	0.7% 19	0.3% 4
Points of interest in the vicinity with routes towards them	0.7% 43	0.7% 17	0.7% 20	0.5% 6
Alternative means of transport towards the destination	3.2% 207	1.7% 40	3.1% 90	6.5% 77

Table 5.11: *Traffic information contents for different motives in % and absolute numbers*

	All situations	Recreational situations	Business situations	Commute situations
Location and length of traffic jams on the complete road network	6.6% 421	6.5% 153	4.8% 112	8.9% 156
Location, length, cause and expected duration of traffic jams on the complete network	19.7% 1261	17.4% 406	16.8% 391	26.6% 464
Advice on the fastest route towards the destination	33.3% 2134	31.3% 730	37.8% 878	30.1% 526
Advice on the shortest route towards the destination	10.9% 699	10.3% 240	11.7% 273	10.6% 186
Advice on the route with the best scenery	3.2% 204	6.7% 156	0.5% 11	2.1% 37
Advice on a parking place and the route towards it	1.3% 86	3.0% 69	0.3% 7	0.6% 10
Expected time of arrival with a margin of 5% for a personal standard route	6.7% 432	6.3% 146	7.3% 170	6.6% 116
Expected time of arrival with a margin of 5% for an advised route	13.6% 871	11.0% 258	18.2% 422	10.9% 191
Gas stations, hotels and restaurants in the vicinity with routes towards them	0.8% 49	1.2% 28	0.4% 9	0.7% 12
Points of interest in the vicinity with routes towards them	0.7% 43	1.2% 28	0.3% 7	0.5% 8
Alternative means of transport towards the destination	3.2% 207	5.2% 122	1.9% 44	2.3% 41

go to areas which they are not very familiar with, and thus want other information or information that provides more comfort to their travel.

Motives and preferences

It was expected that motives used in traffic situation also have an influence on the preference of drivers for traffic information content. In the survey, three motives were used: (a) recreational, (b) business and (c) commute. The results of the survey for these motives are presented in table 5.11. The preference for contents of traffic information differed significantly from the motives according to the Chi-Square test ($p<0.001$).

For the recreational motive, respondents showed a higher preference for other types of information not directly related to travel times. For example, advice on a route with the best scenery or alternative means of transport. Because of the recreational motive, fewer time constraints apply to the trip, thus allowing for other routes or modes. Interestingly, however, the increase found for advice on a parking place, although such advice would perfectly match the recreational motive, is only moderate. The business motive shows clearly that time constraints apply. The respondents preferred both advice on the fastest route and the expected time of arrival for an advised route, as can be expected for such a motive. With a commute motive, most respondents still prefer advice on the fastest route, but the preference for extended traffic information is much greater compared to the other motives. During a commute, drivers probably have more knowledge on the available routes and want to be able to make choices themselves.

Familiarity and preferences

Familiarity was also expected to be of importance for the preference of drivers for contents of traffic information. Situations differed on the hypothetical familiarity of drivers with the road network, but also on the "normal" traffic situation that occurs on this road network. Respondents were considered to be either fully familiar, or not familiar at all. The results of the survey can be found in table 5.12. Again, a significant difference was found among the groups of familiarity according to the Chi-Square test ($p<0.001$).

As was expected, drivers who are considered to be familiar prefer information that enables them to make choices themselves. This results in above-average numbers of respondents choosing the (extended) traffic information. Unfamiliar drivers, however, show the opposite, with higher than average choices for advice on the fastest and shortest route. However, for both groups, the expected time of arrival for an advised route is preferred equally.

Influence of personal characteristics

As with the current usage of traffic information, the relationship between the choice for contents of traffic information with the basic sample characteristics are investigated to find out wether there are certain groups that prefer certain types of traffic information. The analysis starts with the influence of sample characteristics for all situations combined.

For all the situations combined, a significant relationship was found between the attitude towards mobility and usage of traffic information during the trip and the choice for contents of traffic information. The results can be found in tables 5.13 and 5.14.

The big differences between the groups of attitudes towards mobility occur in advice on the fastest route. The Deliberates clearly have the highest preference for this type of information. Strikingly, the Competitors show a relatively high preference for advice on a route with the best scenery, which doesn't match the principles of this group. This group also has the lowest preference for information on alternative modes of transport, which coincides with the individual nature of

Table 5.12: *Traffic information contents for different familiarity in % and absolute numbers*

	All situations	Knowledge of road network	No knowledge of road network
Location and length of traffic jams on the complete road network	6.6% 421	7.6% 310	4.8% 111
Location, length, cause and expected duration of traffic jams on the complete network	19.7% 1261	23.0% 940	13.8% 321
Advice on the fastest route towards the destination	33.3% 2134	31.6% 1293	36.3% 841
Advice on the shortest route towards the destination	10.9% 699	9.6% 394	13.2% 305
Advice on the route with the best scenery	3.2% 204	1.9% 76	5.5% 128
Advice on a parking place and the route towards it	1.3% 86	1.7% 69	0.7% 17
Expected time of arrival with a margin of 5% for a personal standard route	6.7% 432	6.8% 279	6.6% 153
Expected time of arrival with a margin of 5% for an advised route	13.6% 871	13.5% 551	13.8% 320
Gas stations, hotels and restaurants in the vicinity with routes towards them	0.8% 49	0.5% 20	1.3% 29
Points of interest in the vicinity with routes towards them	0.7% 43	0.4% 16	1.2% 27
Alternative means of transport towards the destination	3.2% 207	3.4% 141	2.8% 66

Table 5.13: *Choice for information contents vs Attitude towards mobility*

	Accepter	**Deliberate**	**Competitor**	**Conscious**	**Enjoyers**
Location and length of traffic jams on the complete road network	9.9%	10.4%	9.3%	9.7%	7.2%
Location, length, cause and expected duration of traffic jams on the complete network	16.2%	19.3%	18.9%	16.7%	16.3%
Advice for the fastest route towards the destination	22.6%	25.7%	21.8%	18.8%	19.9%
Advice for the shortest route towards the destination	12.6%	11.0%	13.1%	13.9%	14.4%
Advice for the route with the best scenery	3.4%	2.4%	8.3%	8.7%	7.8%
Advice for a parking place and the route towards it	3.4%	2.0%	3.5%	1.0%	3.6%
Expected time of arrival with a margin of 5% for a personal standard route	8.9%	8.6%	7.7%	7.6%	10.1%
Expected time of arrival with a margin of 5% for an advised route	12.8%	13.7%	11.5%	12.5%	13.4%
Gas stations, hotels and restaurants in the vicinity with routes towards them	1.7%	0.8%	1.6%	2.4%	0.7%
Points of interest in the vicinity with routes towards them	2.2%	0.2%	1.6%	2.1%	1.3%
Alternative means of transport towards the destination	6.2%	6.0%	2.6%	6.6%	5.2%

Table 5.14: *Choice for information contents vs usage of traffic information during travel*

	Are you using traffic information during your trip?	
	Yes	**No**
Location and length of traffic jams on the complete road network	9.5%	9.6%
Location, length, cause and expected duration of traffic jams on the complete network	18.3%	15.6%
Advice for the fastest route towards the destination	22.1%	22.1%
Advice for the shortest route towards the destination	12.1%	13.9%
Advice for the route with the best scenery	5.6%	5.5%
Advice for a parking place and the route towards it	2.9%	2.6%
Expected time of arrival with a margin of 5% for a personal standard route	8.7%	9.2%
Expected time of arrival with a margin of 5% for an advised route	13.8%	10.2%
Gas stations, hotels and restaurants in the vicinity with routes towards them	1.1%	2.3%
Points of interest in the vicinity with routes towards them	1.0%	2.6%
Alternative means of transport towards the destination	4.9%	6.0%

this group. Also the Conscious and Enjoyer have a large preference for a route with the best scenery, which can be explained by the background of both groups.

When taking into account whether respondents use traffic information during their trips, it becomes clear that those who are using traffic information have different preferences, especially regarding the congestion information. When one uses traffic information during their travel, they prefer the extended congestion information, significantly more than those who do not use traffic information. This also applies to the expected time of arrival for an advised route. In contrast, respondents not using traffic information have a higher preference for an advice on the shortest route. It seems as if the respondent who uses traffic information is more demanding in the preferred content of the traffic information.

Based on analysis of all the types of traffic situation (non-recurrent, various motives and familiarity) three characteristics seem to be the most important in relation to the choice of contents for traffic information. These three are:

- attitude towards mobility;
- age;
- usage of traffic information during travel.

Table 5.15: *Types of information*

Alternative	Type of traffic information	#	%
1	Informing information	705	13.6%
2	Advising information	860	16.6%
3	Personalized information	872	16,9%
4	Predictive information	693	13.4%
5	No information	2045	39.5%

The results indicate that the groups defined by their attitude towards mobility all have significant differences in respect of specific contents of traffic information, usually related to the background of these groups. For example, a respondent in the conscious group is more likely to choose advice on the route with the best scenery, while a respondent in the deliberate group is more likely to choose advice on the fastest route. For age, one major item was found. The older a respondent, the more likely he is to choose advice on the shortest route. This is most likely caused by a misinterpretation of the shortest route, because the older people expect the shortest route to be the same as the fastest. As already mentioned, respondents who are using traffic information during their travel have more demands regarding the content of traffic information, especially when it comes down to expected time of arrival or congestion information.

5.3.5 Traffic information attributes

As described in the approach (section 5.2), the attributes of information were also used in the survey, using a stated choice approach. In this approach, five types of traffic information were used, in contrast to the previous section, where eleven different contents for traffic information were used. Respondents made nine choices each, resulting in a total number of 5175 choices. The distribution of these choices over the alternatives can be found in table 5.15. It is obvious that most respondents chose to have “no information”, probably caused by the costs that were attributes of the other alternatives. Of the four other alternatives, the “advising information” and the “personalized information” show equal choices, as do the “informing information” and the “predictive information”. These differences are probably caused by the fact that “informing information” is not much different from the currently available information, except for costs. “Predictive information” probably shows a smaller amount of choices because of its predictive nature. Respondents might not trust these predictions.

With discrete choice estimation software Biogeme, several models were estimated. The starting point here is the multinomial logit model, given in equation (4.5). The multinomial logit model was first applied to the equation given in equation (5.1) and again repeated for convenience.

$$\begin{aligned} V_i =& ASC_i + \beta_{cost} Cost_i + \beta_{timeliness} Timeliness_i \\ &+ \beta_{reliability} Reliability_i \end{aligned} \tag{5.4}$$

In the first application of the multinomial logit model (MNL1) the coefficients for the attributes of the information were dropped, resulting in a model that only

Table 5.16: *Estimated MNL models*

Variable	MNL1		MNL2	
	Estimate	t statistic	Estimate	t statistic
ASC1 (informing)	0		0	
ASC2 (advising)	0.198	3.91	0.211	3.72
ASC3 (personalized)	0.212	4.20	0.227	4.02
ASC4 (predictive)	−0.017[a]	−0.32[a]	−0.039[a]	−0.67[a]
ASC5 (none)	1.065	24.38	0.240	4.44
β_{cost}			−2.339	−35.80
$\beta_{timeliness}$			−0.481	−18.16
$\beta_{reliability}$			0.988	33.71
Number of observations	5175		5175	
Null log likelihood	−8328.84		−8328.84	
Final log likelihood	−7793.6		−6241.61	
Likelihood ratio test	1070.48		4174.47	
Pseudo ρ^2	0.0643		0.2506	
Adjusted pseudo ρ^2	0.0637		0.2498	

[a] Non-significant estimate at a 95% level.

captures the average preference for a certain type, with the alternative specific constant. The second model (MNL2) fully used equation (5.1) for all alternatives apart from the "no information" alternative (which has no attributes). The results of the estimation are presented in table 5.16. The values of the alternative specific constant are valued relative to the base alternative, which was chosen as the first alternative, "informing information".

The first model, only using the alternative specific constants, does not show a good model fit according to the pseudo ρ^2 although the likelihood ratio test indicates that all parameters jointly differ from zero (Chi-Square, $p<0.05$). It is clear that alternative 5, "no information" is the most preferred option relative to the others, as it has the highest positive value. Alternatives 2 "advising information" and 3 "personalized information" also have a positive value, which indicates that these two alternatives are also preferred to the base alternative. For all these alternatives, the t-test also shows that the values that are estimated are significant. The preference for the "no information" alternative is easily explained. First of all, traffic information is considered to be freely available. Drivers do not pay for receiving information via radio bulletins. Second, the survey showed that most respondents were satisfied or very satisfied with the information they used, so no improvements seem necessary.

The second model shows a much better fit, as indicated by the higher pseudo ρ^2. The likelihood ratio test again indicates that all parameters jointly differ from zero (Chi-Square, $p<0.05$). All the estimated coefficients have the expected sign. Again, the alternative "no information" is preferred to the other alternatives; however, the differences are much smaller compared to the first model. The alternative specific constant of alternative 4 (predictive information) is not significantly different from zero according to the t statistic, indicating no different preference compared to the informing information. The difference in alternative specific constant for alternative 5, between both models, is striking, caused by the existence of the extra information carried in the estimated attribute coefficients.

Table 5.17: *Estimated MNL models for all attitudes toward mobility*

Variable	Accepters	Deliberates	Conscious	Competitors	Enjoyers
ASC1 (informing)	0	0	0	0	0
ASC2 (advising)	0.35	0.03[a]	0.18[a]	0.19[a]	0.30[a]
ASC3 (personalized)	0.31	0.07[a]	0.11[a]	0.36	0.35
ASC4 (predictive)	0.03[a]	−0.18[a]	−0.20[a]	−0.01[a]	0.13[a]
ASC5 (none)	0.21	−0.10[a]	0.12[a]	0.68	0.40
β_{cost}	−2.33	−2.84	−2.34	−1.82	−2.00
$\beta_{timeliness}$	−0.42	−0.55	−0.51	−0.54	−0.41
$\beta_{reliability}$	0.95	1.06	0.82	1.04	1.02
Number of observations	1454	1346	718	652	647
Null log likelihood	−2340.12	−2166.30	−1155.58	−1049.35	−1041.31
Final log likelihood	−1787.38	−1539.16	−884.54	−787.29	−804.66
Likelihood ratio test	1105.48	1254.28	542.08	524.12	473.29
Pseudo ρ^2	0.236	0.289	0.235	0.250	0.227
Adjusted pseudo ρ^2	0.233	0.286	0.228	0.243	0.220

[a] Non-significant estimate at a 95% level.

The coefficients for costs (β_{cost}) show a strong negative relationship. When the costs of an alternative increase, the number of choices for this alternative decreases quickly. This again concurs with the fact that traffic information is freely available at the moment. Drivers obviously do not want to pay (much) for their information. The coefficients for timeliness ($\beta_{timeliness}$) also show a negative relationship, in line with the coding used for the levels. When information is provided in real time the utility of the alternative will increase. For large delays, the utility will decrease. Timeliness, when compared to costs, is valued less negative by the respondents. The coefficients for reliability ($\beta_{reliability}$) also show signs matching expectations. If the reliability is increased (a positive level), the utility increases and vice versa. The size of the coefficients shows that reliability is valued more than timeliness. The overall importance of the characteristics thus shows that costs are the most important, followed by reliability and timeliness.

Attitude towards mobility

One of the extra explanatory variables, also used in the study of the contents of traffic information, is the mobility type. In the case of multinomial logit models, several options are available to incorporate such a variable in the model. One of those involves using dummy variables, each which has its own β. In this case, however, separate models were estimated for each mobility type. In essence, this provides the same results. The estimated parameters are presented in table 5.17.

First of all, all the different models show more or less equal fits, and the likelihood ratio test for all models show that the estimated parameters jointly differ from zero. However, when looking at the specific estimates, a number of those are not significantly different from zero (these values are emphasized in the table).

The estimated variables also show some interesting differences between the attitudes toward mobility. First of all, only the accepters have a significant value for the alternative specific constant of advising information, meaning that the accepters prefer advising information to informing information. Personalized information is significantly preferred to informing information again by the accepters, but also by the competitors and enjoyers. This seems reasonable when looking at the description of these mobility types. Enjoyers, for example, see mobility as fun, as a way to socialize. This could mean that travelling is not bad, and that they want to be able to have nice routes, which is possible if the information is personalized. More or less the same applies to competitors, where the need to be better than others is important, also indicating a need for personalized information. Predictive information is never significantly preferred over informing information. Accepters, competitors and enjoyers also show a significant preference for no information, compared to informing information. This means that they would rather not pay for information and thus won't have any information available. The mobility types might thus be more sensitive to costs than other types.

For the estimated β's too some interesting differences are found between the attitudes toward mobility. Regarding costs, the competitors show the least negative value, meaning that costs are the least important to this group compared to others. Deliberates, on the other hand, have a much higher negative value, which means that cost is very important for this group. Timeliness is most important for the deliberates, conscious and competitors, because these groups have the most negative estimates (which is in line with the coding used, where real-time information has a negative coding). Real time information is thus appreciated most by these groups, which is easily explained for the competitors because they want to be faster than others, which is easier if the information available is current. More or less the same applies to the deliberates who want to minimize their time spent in traffic. Reliability is more or less equally important to all groups, apart from the conscious who show a slightly lower value. Reliability is not as important to them as it is to other groups. This could coincide with the fact that conscious travellers are more prone to use public transport, which is said to be less reliable than car traffic.

Similar alternatives

The MNL models as estimated in the previous section all use the assumption of independent and identically distributed errors. However, except for the no information alternative, the alternatives are likely to violate this assumption which is explained in section 4.2.2. The alternatives concerning information are likely to differ from the no information alternative. Accommodating this violation of the i.i.d assumption is possible using the nested logit model. In this approach, the four alternatives with information are expected to be correlated, within a partition (the information nest), but not with the no information alternative (the other nest). The resulting estimates given in table 5.18 are based on a partially degenerate nested logit model given in figure 5.3.

A first look at the results show that the estimated parameter signs are all in accordance with expectations, which gives a first validation of the estimated model.

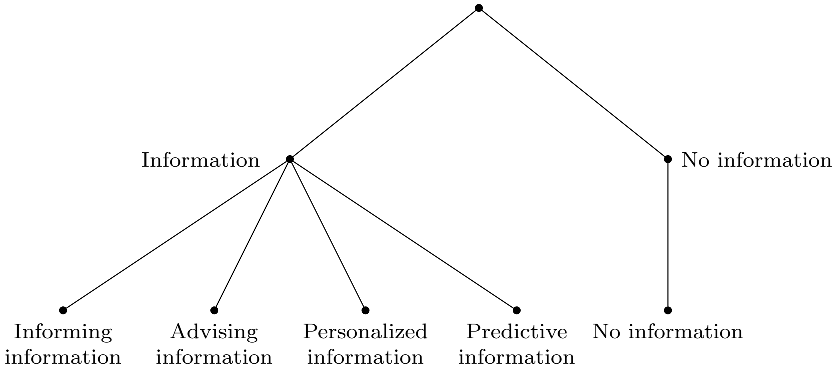


Figure 5.3: *Nest structure used for NL*

Table 5.18: *Estimated NL model (using RU2)*

Variable	**NL** Estimate	*t* statistic
ASC1 (informing)	0	
ASC2 (advising)	0.136	4.24
ASC3 (personalized)	0.125	3.82
ASC4 (predictive)	−0.036[a]	−1.10[a]
ASC5 (none)	−0.103	−2.28
β_{cost}	−1.377	−14.04
$\beta_{timeliness}$	−0.285	−11.98
$\beta_{reliability}$	0.595	14.15
NestA (information alternatives)	1.971	(0) 13.56
		(1) 6.68
NestB (no information alternative)	1	
Number of observations	5175	
Null log likelihood	−8328.84	
Final log likelihood	−6179.3	
Likelihood ratio test	4299.19	
Pseudo ρ^2	0.2581	
Adjusted pseudo ρ^2	0.2571	

[a] Non-significant estimate at a 95% level.

A better look shows that the parameters are jointly different from zero (Chi-Square, $p<0.05$). The model fit, in terms of the pseudo ρ^2 is approximately as large as in the MNL2 model. However, another likelihood ratio test to compare the final log likelihoods of the MNL2 and NL model indicates that there is a significant difference in both models. The nested logit model thus provides a significantly better estimate.

The big difference in the estimates is caused by the nest parameter. As described in section 4.2.3, this parameter describes the correlation of the alternatives within a nest, when used in equation (4.13). In this case, because of the applied normalization (which is RU2, or normalization from the top), the nest parameter estimated here is λ_m. This results in a correlation of 0.74, which indicates a pretty strong correlation of the alternatives in nest A. The t-statistic for the nest parameter also shows that this parameter is significantly different from both zero and one. This indicates that the nested logit model provides a better estimate than the multinomial logit model.

The parameter estimates for all the β's are as expected (see table 5.18). Compared to the previously estimated MNL2 model, the size of the parameters is somewhat smaller. This is, of course, caused by the introduction of the two nests and their scale parameters. Interestingly, the results for the alternative specific constants do show a bigger difference, especially for the "no information" alternative. This alternative is no longer preferred to the "informing information", compared to the MNL2 model. This indicates that "no information" is not really an option for the respondents. However, the negative utility of costs of the information alternatives eventually makes them choose no traffic information. The "predictive information" alternative again shows no significant difference from the "informing information" alternative, while the other two alternatives more or less have the same preference to this "informing information" alternative.

Mixed panel logit

The next estimated model concerns the panel version of mixed logit, as explained in section 4.2.5 and equation (4.20). In the survey, a participant made a series of nine consecutive choices. This series of choices violates the i.i.d. assumption of the MNL model, as the same respondent is likely to be influenced by previous choices. The same respondent might favor one of the alternatives, and thus show correlation in the choices in one series. Accommodating this violation is possible using the panel mixed logit model. In this model it was chosen to use the normal distribution, where the average is given by the alternative specific constant, with a variance added. With 1000 Halton draws, the results are as given in table 5.19. Each alternative specific constant has its own associated variance, which allows for alternative specific variations between people in the panel, except for the alternative that is given a constant for the ASC. The variances are given by σ in table 5.19.

The results show an enormous increase in model fit, compared to the standard multinomial logit model as provided in table 5.16. The Pseudo ρ^2 increases from 0.25 to 0.45, almost double. The likelihood ratio test also shows that the panel mixed logit model is significantly better than the multinomial and the nested logit models. The inclusion of taste variation by allowing for heterogeneity in the alternative specific constants thus greatly improves the model fit. Respondents clearly differed in their preference for both the alternatives and the parameters, simultaneously and separately.

The parameter estimates are all in accordance with the expected values, but they do show some differences to the previously estimated models, especially for the alternative specific constant of the "no information" alternative. In the multinomial logit model it was significantly preferred over the "informing information" alternative. In this case, it is significantly different, in the sense that is valued less than the "informing information". However, the variance of this estimate is significantly different from zero, and has the largest value of all estimated variances. This shows that there is a large taste variation among the respondents who chose this alternative, of which around half have a positive preference, and the other half have a negative preference for having no information available at all.

Table 5.19: *Estimated Panel Mixed Logit model*

Variable	**Panel Mixed Logit** Estimate	*t* statistic
ASC1 (informing)	0	
σ_1 (informing)	0	
ASC2 (advising)	0.17	2.23
σ_2 (advising)	−0.57	−5.01
ASC3 (personalized)	0.10[a]	1.22[a]
σ_3 (personalized)	0.61	−6.72
ASC4 (predictive)	−0.11[a]	−1.31[a]
σ_4 (predictive)	0.91	9.68
ASC5 (none)	−0.45	−2.51
σ_5 (none)	5.19	17.51
β_{cost}	−3.21	−36.01
$\beta_{timeliness}$	−0.66	−19.96
$\beta_{reliability}$	1.38	33.55
Number of observations	5175	
Null log likelihood	−8328.84	
Final log likelihood	−4502.59	
Likelihood ratio test	7652.49	
Pseudo ρ^2	0.4594	
Adjusted pseudo ρ^2	0.4581	

[a] Non-significant estimate at a 95% level.

Not surprisingly, the variance of the other alternatives is much lower without being insignificant, which indicates that although taste variations exist among the respondents, the variations are not large at all. The alternative with the highest variance is "predictive information" which can be explained by the trust of the respondents in prediction. One who believes that predictions provide reasonable and reliable results is much more likely to choose this alternative, whereas one who doesn't believe in predictions is biased against this alternative. Because no parameter for trust is included, the variance for this alternative is higher.

The estimated values for the attribute parameters all are in accordance with expectation, regarding their direction and size. The most important attribute remains the cost, and compared to the previous models, the size is larger. Taking into account taste variations around a mean, in the form of an alternative specific constant, this shows an increase in sizes of the attribute values, which is in extra indication that there is at least some taste variation in the population. Costs thus became more important to respondents. The same applies to the reliability and the timeliness of the traffic information.

Also a combination of the nested logit and mixed panel logit has been applied to the dataset. These results however did not provide an improvement in the estimation. Besides, the estimation did not find a significant difference between both nests. Therefore, the mixed panel logit is deemed to have the best fit, while accommodating for the possible measurement effects using the sequence of choices.

5.4 Discussion

5.4.1 Survey, model and data integrity

Internet and respondents

The results in this report are based on an Internet survey. Apart from the many benefits of such a survey, there is also one important drawback. Access of the survey through internet introduces a bias because only people who have a connection to the internet can fill in the survey. According to the Bureau for Statistics in the Netherlands, 83% of the Dutch population (between 12 and 75 years old) is able to access internet (CBS, 2011). Of this group, 54% uses internet on a daily basis, for various purposes. Although these figures indicate that a large group of the Dutch population could participate in this survey, the results are probably still biased because important groups (people older than 65, and people with a lower education) have a much lower percentage of internet access. These two groups are also under-represented in the sample of the survey.

Another drawback exists, namely the way in which respondents were asked to participate. The self-gathered group consists of personal and business contacts which can lead to a bias because of their familiarity with the subject of study. The other group was invited via an agency which provided an internet panel. Both groups belong to the target group because of their possession of a driving license, which was checked within the survey. They both faced the same questions in the survey. However, for the internet panel it was possible to specify the characteristics of the sample in terms of the average Dutch possessor of a driving license. This of course was not possible for the personal and business contacts. This could have introduced a difference in the results between both groups of respondents.

Both groups of respondents were first compared on the basic sample characteristics, with the Chi-Square test. Results of this test show that only the most used motive for car use did not differ significantly between both groups whereas all other characteristics (age class, sex, level of education, mobility type, household size and kilometers per year) are not the same in both groups. For example, the internet panel consisted of more females, more older age groups, less highly educated respondents, more large households, and more respondents with fewer kilometers per year. The same kind of tests were also performed on the information usage. In this case possession of a navigation system and use of traffic information before travel were not significantly different between both groups. Using information during travel and satisfaction with the information used is different. For example, the internet panel has more respondents not using information during travel and in general are more satisfied with the traffic information used. When looking at the choice of information contents, a combination of all situations is used. In this case again a significant difference is found between both sources for advice on the fastest and shortest route, as well as the estimated time of arrival for an advised route. The internet panel has fewer respondents who choose the fastest route and the estimated time of arrival, but has more respondents choosing the shortest route. The differences that occur between both groups are expected to be caused by the actual respondents in each group. This would indicate a difference in basic sample characteristics (age, gender, level of education).

Survey design

Apart from these drawbacks of the internet survey, other, more general causes for bias may exist in the survey. First of all, it is not known to what extent the respondents completely understood all the different situations (eleven in total) and the different options for contents of the traffic information (the fourth part of the survey). The description of the situations were quite precise, with the important characteristics of the situation in bold. This also caused the description to be rather long, especially if eleven situations need to be described. This way of describing was based on the remarks received after pilot testing the survey. For the options the respondents could choose from, no further description was provided. This may have caused some problems with respondents not understanding the difference between the shortest and the fastest route. Randomization of the options per situation at least guaranteed that respondents had to look at all the options to find their preferred one for each situation.

Of course, some elements that can influence the choice of respondents were not part of the survey. Examples are their experience with congestion and more questions on their experiences with traffic information. This would have meant more personalization of the survey, which would make the analysis more difficult. Such elements need to be taken into account with further study in the route choice simulator (see also chapter 7).

Choice models

As already mentioned in the approach (see section 5.2) it is assumed in the multinomial logit model that the choices of all respondents are independent. However, because one respondent has to make nine choices, this assumption is violated. The results of the MNL model therefore have to be used with great care. Another assumption is that the alternatives are independent of each other. This assumption can also be questioned when looking at the alternatives. Four alternatives present information, while the fifth alternative does not. This indicates a relationship between the first four alternatives, which violates the assumption of independence. Both violated assumptions were relieved by using other models, such as nested logit and mixed logit for panel data.

The introduction of these more advanced models causes other points of discussion. First of all, the nested logit model that was applied still assumes independent choices of the respondents, which is violated. It does allow for similarity between the alternatives, by putting the relevant alternatives in one nest. In this study, not all possible variants of the nests were used. The model that performs best does not necessarily have to be based on assumptions of similarity made by the researcher.

The mixed logit model presents a number of great advances in modelling stated choices, but these advances also introduce a number of problems. First of all, the researcher has to define which of the possible approaches is used: (a) error components or (b) random coefficients. In essence, the results of the approaches are the same, but the way one looks at the utility function differs. The second issue is the selection of the random parameters or, in other words, the attributes which are allowed to have a mean and variance. This also introduces the third problem, the selection of the distribution for this parameter. Each distribution has its own implications. For example, a Normal distribution has long tails on each side, which implicates that some individuals in the estimated distribution will either have an extremely high or low value given for that attribute. It is clear that this is not always realistic. Other distributions do not show this problem, and can be tuned to not allow negative values, such as the triangular distribution. The Normal distribution was used in this research because of time constraints in the modelling. In further research, the results of different distributions should be investigated.

The fourth issue is to take into account the correlation between choice situations. In this research the approach was taken to use a random coefficient for the alternative specific constant, but other options are also available, which might have an influence on the outcome of the estimation. The fifth issue discussed here is the number of draws used in the estimation. In this case it was chosen to use 1000 draws, based on recommendations in the literature.

Another, last issue, not taken into account in this research, is the correlation between the parameters. Several options are available here. One can use interaction, where a parameter is estimated for two (or more) multiplied attributes. This has effects on the number of questions necessary for the survey, which increases fast. Another option, introduced by the mixed logit model, is to change the variance-covariance matrix. This is a very complex approach, but can provide very good results. However, because of the complexity, this has not been undertaken in this research.

In the research for the attributes of traffic information, it is also possible to use more socio-demographic attributes than are currently used. Separate models have been estimated for the different attitudes toward mobility, but also age, level of education etc. can be used here. Several options are available here. First, one can estimate separate models for each group, as was done for the attitude towards mobility. Second, one can use different scale parameters for each group. For more than one socio-demographic attribute, this would mean a large number of scale parameters. The third option, which is the preferred option, is to introduce dummy variables for each attribute level. For an attribute with two levels (for example, wether one owns a car or not) this would mean one dummy, either zero (no vehicle) or one (one or more vehicles). For a three level attribute, two dummies are necessary, etc. This approach means that the data that is used has to be redefined, to allow for these dummy variables to be measured. Since this procedure is very time consuming, the dummies have not been created, and hence not estimated. In future research, the creation of dummy variables in the data should be done earlier on.

5.4.2 Results

For the results of the survey for the contents of traffic information all results make sense. As was expected, the situations are of significant influence on the choice of contents of traffic information, although the most preferred option stays the same over the different alternatives. The advice for the fastest route that was most preferred is also a logical finding. Overall, one can say that the results from the survey are logical.

Compared to the presented literature, the results of this survey are in line with previous findings. The basic preferences of drivers are with traffic information related to travel times, either directly in the form of an advice or estimated time of arrival, or indirectly by receiving detailed information on location, length, duration and cause of traffic jams. Interestingly, however, the percentage of respondents interested in advice is higher than could be expected from previous studies, where the extended information seems to be preferred (see for instance the results of (Mensonides, 2004)). Another important issue here is that the results provide more insight into differences in information preferences between different traffic situations, either with a different non-recurrence, motive or familiarity. The findings regarding personal characteristics in the presented results also are in line with findings in previous research. For example, the attitude towards mobility is an important factor and is of significant influence on the choices of information types and contents of traffic information. Also familiarity and motive are important factors, which were also found in other studies.

The attributes of traffic information, as presented in section 5.3.4 are overall as was found in previous studies. The attributes used here are all significant in choosing a type of information, with costs being the most important. However, in a previous study, accuracy was found to be more important than costs (Ng et al., 1995). However, from the study it is not clear how accuracy was defined, which means that there could be some overlap between reliability in this study and the accuracy defined in the study of Ng et al. (1995).

60% of the respondents is willing to pay for a variant of traffic information, with varying levels of reliability and timeliness and 40% of the respondents chose not to have any information. This is a very high percentage of travellers paying for traffic information. It is expected that this value is an overestimation of the actual willingness to pay, since in the survey there was no actual cost involved for the respondents, reducing the reliability of such an survey. However, without costs as an attribute, each respondent would expected to be always choosing the best option. With costs, a trade-off must be made between costs and other attributes.

5.5 Summary

In this chapter the focus was on the results of the survey as designed in section 5.2 that presented a conceptual model, which led to a number of hypotheses regarding the preferences for traffic information. First of all, the relationship between information contents and situations was investigated. This showed that the situation a driver is facing influences the choice of traffic information contents. The influential characteristics of the situation are the actual traffic situation at hand (accidents, road works and large events), the motive a driver has (commute, business or recreational) and the familiarity of the driver with the road network and traffic situations. The most preferred content of traffic information for all situations is advice on the fastest route, followed by extended traffic information and expected time of arrival. This clearly shows that irrespective of the situation, drivers prefer information that concerns travel time or is related to that. Depending on the actual situation, other types of information are chosen more often, for example information concerning comfort.

The personal characteristics, together with the situational characteristics were also investigated in relation to the choice of information contents. Based on the analysis, three main personal characteristics are important, irrespective of the situation:

- attitude towards mobility;
- age;
- usage of traffic information during travel.

Partially related to the contents of information, the stated preference survey investigated the choice of types of traffic information together with the important attributes of these types. Based on the total choices made, the driver is most interested in advice or even a personal advice. The predictive and informing information are a little less preferred by the respondents. A total of 60% of the respondents were willing to pay for traffic information whereas the remaining 40% choose not to have traffic information. This striking fact is due to the costs involved in the other types of traffic information. The fact that advice was most often picked, relates to the previously found preferred content for traffic information, which also mentioned the advice on a fastest route.

The data of the survey for information attributes was used to estimate several discrete choice models: multinomial logit, nested logit and mixed logit for panel data. The results of these models indicate that the most important attribute for traffic information is the cost. Increasing cost deters a respondent from using an alternative. Next most important is the reliability, followed by timeliness. This order of importance for the attributes stays the same in all applied models. The multinomial logit model was also used for the different attitudes towards mobility. The results indicate that there are differences in valuation of the attributes and alternatives between the groups, which is in accordance with the results of the relationship between contents of traffic information and attitude towards mobility.

The nested logit model showed that there is a significant correlation between the alternatives that provide information, which means that they have something in common, which of course is in accordance with the description of the alternatives.

The four alternatives in which there is information available can be grouped together, with a correlation of 0.74. The last applied model, which is based on mixed logit for panel data, showed that there is great variation between the respondents for the attributes when looking at the alternatives available, especially the variations in taste for the "no information" alternative that were very large. Of all the models that were applied to the dataset, the mixed logit model for panel data performed the best in terms of the pseudo ρ^2 and likelihood ratio test.

Based on these results, it is clear that the preference for traffic information changes according to situations, motives, knowledge and drivers. The choice for information also depends on its attributes, especially costs. It indicates a need for personalization of traffic information which is specific to the different preferences and situations. The results are used in designing an experiment for route choice behaviour during non-recurrent traffic situations (see chapter 7).

Errors using inadequate data are much less than those using no data at all.

— Charles Babbage

Chapter 6

Route choice data collection

6.1 Introduction

This chapter presents the collection of data for the estimation of route choices models to be estimated (based on discrete choice). It provides a partial answer to research question three, especially part a: *How can route choice behaviour be measured?* Measuring route choice is possible with various methods, each with its advantages and disadvantages. For this research, data is needed to estimate a route choice model that indicates the behaviour of car drivers during non-recurrent traffic situations in the presence of traffic information. Based on a literature review a choice is made to use a route choice simulator. Several items are of importance for the design of a route choice simulator, especially for the non-recurrent (accident) traffic situations. These items are all discussed and dealt with in the design of the route choice simulator in this chapter.

The chapter starts with a discussion of the possibilities for collection choice data in general, and route choice data in particular (section 6.2). We will discuss the possibilities and suitability of revealed and state preference for this research. The choice for a route simulator as having the advantages of both stated and revealed preference is further explained with its design in section 6.3. For route choice simulators (RCS) (and stated preference data in general) it is important to investigate the validity of the resulting data, models and coefficients that are estimated using this data. For route choice simulators and the AIDA RCS, validity is discussed in section 6.4. The chapter finishes with a summary in section 6.5.

6.2 Route choice data types and sources

6.2.1 Data types

Choice or preference data in general come in many forms and can have many sources. The data can generally be classified in three types:

- data generated using models;
- data of what people actually do (revealed), and;
- data of what people intend or say they will do (stated).

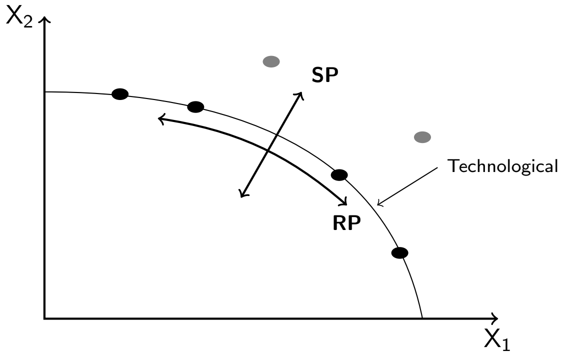


Figure 6.1: *The technological frontier and Revealed and Stated Preference data (Louviere et al., 2000)*

The first of these types concerns a specific type of data compared to the other two. This so called synthetic data is generated using a predetermined model and is most often applied to present developments or variations in modelling approaches. For example, when applying different hypotheses to a discrete choice model which assume heteroscedasticity and/or heterogeneity, it is very helpful to create a synthetic dataset for which you know (because you have set them at specific values) the correct estimates. Because of these known estimates, various modelling approaches can be compared not just based on model fit, but also in difference between the actual and estimated parameter values. An example of such an approach can be found in Munizagaa et al. (2000); Hess (2005). Another type of application concerns route choice modelling. In case of overlap in routes, a simple hypothetical network can be used to explain the different approaches that deal with this overlap (see for example Ramming (2001); Frejinger (2008); Bovy et al. (2008); Fosgerau et al. (2009)).

The other two types of data are called revealed and stated preference or choice. Revealed preference (RP) data concerns events or choices that have been observed to have actually occurred. For route choice it would for example be a sequence of choices at all junctions a driver passes in a trip. Each choice is observed. Stated preference (SP) data is about choices made in hypothetical situations. It is not an observation of a choice that actually occurred, but an observation of the intention to make such a choice when facing the same situation in reality. Revealed choice data can only represent data on choice made in actual situations, whereas stated choice data allows for both actual and non-existing situations. This can also be understood using figure 6.1, which shows two random attributes X_1 and X_2 with different levels. This figure indicates the so called technological frontier which makes the difference between actual market situations that can be used for observations of choices, and the not (yet) existing market situations where no actual choices can be observed, simply because the situations or alternatives do not exist. Revealed choice can only include existing alternatives with know attributes and attribute values. This relates to the actual market situation. Stated preference can also include new alternatives and other or new attributes which cannot be found in the market, and as such crosses this technological frontier.

Table 6.1: *Summary of RP and SP data (Louviere et al., 2000)*

RP	SP
• Depict the world as it is now • Possess inherent relationships between attributes • Have only existing alternatives as observables • Embody market and personal constraints on the decision maker • Have high reliability and face validity • Yield one observation per respondent at each observation point	• Describe hypothetical or virtual decision contexts • Control relationships between attributes which permits mapping of utility functions with technologies different from existing ones • Can include existing and/or proposed and/or generic choice alternatives • Cannot easily represent changes in market and personal constraints effectively • Seem to be reliable when respondents understand, are committed to and can respond to tasks • Quite often yield multiple observations per respondent at each observation point

A summary of characteristics of RP and SP data is given by Louviere et al. (2000) and reproduced in table 6.1.

Stated preference research allows for much more controllability, because the experiment is designed with a specific context and alternatives with their attributes are provided by the researcher. This enables a researcher to design an experiment which targets the research objectives that are of interest. A revealed preference approach lacks this controllability, because the research is (normally) not able to design the situations and alternatives in the choice situations. This means that although a choice itself is a valid choice, it is not necessarily a choice of interest for the research, as the alternatives or context does not suit the research objectives. Next to that, the revealed preference approach has no influence on or control of the levels of knowledge of individuals, which makes it difficult to know the participants choice set (as just the chosen alternative is observed).

The amount of observations per respondent is a big difference between RP and SP approaches. As for RP data, just a single observation at each point can be found, it is a time consuming (and therefore costly) approach to get enough variability in the choice data. SP data on the other hand allows for multiple observations at each point, even from one participant. It is easier to get the amount of data necessary with less participants. This makes the SP approach less time consuming and more cost effective.

The biggest disadvantage of SP data is the validity. SP research makes use of a situation in which respondents are placed by the researcher, it is important they completely understand this hypothetical context and the tasks they have to perform (Carson et al., 1994). Revealed preference approaches do not have this disadvantage, as the choice is actually made. The participant does not have to imagine a certain situation and think which choice would be made in that situation. SP data can (and perhaps should be) interpreted as a stated behavioural intention (Fujii & Gärling, 2003). Habits, for example, can overrule the stated intention. As mentioned by Bogers (2009) two important issues with validity concern the (usually) lacking of consequences of choices made and the difficulty to convey certain attributes of alternatives such as reliability. In reality, a choice will have a certain consequence in terms of travel time and costs. In a hypothetical context, such a consequence is difficult to implement. It is possible for a researcher to reward a subject with money for good choices, and penalize subject for bad choices (or late arrivals), but by doing so, the researcher influences the choice behaviour of the participant. This can have a significant impact on the validity of the data (Carson & Groves, 2007), as in most cases, the structure and amount of incentives for the choice situation are not known. It is likely that the incentive structure influences choice behaviour in such a way that the resulting data is not valid. However, using an incentive can also increase the realism of the choice situation, which provides for more valid results. The other issue concerns vague or difficult to grasp concepts such as reliability. The way individuals interpret such a concept influences the choices they will state, but as this interpretation is not known, the actual behavioural rule for such a choice is not known. Next, if a concept is not used to describe alternatives and attributes in every day life (such as the reliability of travel time, which is almost never provided as part of traffic information), people do not have a common experience with such a concept. This makes it less likely that they use this information in their actual choice.

Even though the validity of SP research is a real issue, it has seen many applications, as the advantages of this approach provide the possibilities for scientists and analysts to investigate new concepts and alternatives, or situations which would be impossible, too costly or time consuming to study using RP research.

6.2.2 Data sources

In case of route choice, the acquisition of the data is an important element to address. For both SP and RP data, several options are available to acquire the necessary data for route choice modelling. An absolute minimum for the data is to have a selected alternative. Attributes of this alternative, as well as the choice set, can be generated using simulation and modelling. However, the more data is available on the chosen alternative and the choice set, the more valid and better the model will be.

Observing route choice in the real world can be done using several methods. They can be grouped into using surveys and tracking vehicles, where methods can be combined to improve data quality and add more information. The tracking methods are based on observing vehicles and their routes in a road network. This tracking can be done using GPS data (where a GPS devices is located in the vehicle), GSM or bluetooth data (from smartphones) or license plate recognition. Using GPS, GSM or bluetooth data involves a lot of processing before it can be used as route choice data. A thorough examination of this subject can be found in Schüssler (2010). Another approach is to use the concept of domain of data relevance, which corresponds to a physical area in the road network where measurement is relevant (Bierlaire & Frejinger, 2008; Bierlaire et al., 2010; Frejinger, 2008). Using license plates to reconstruct routes is also a possibility, but has other issues. This method only uses a few (depending on the density for observations) locations where a license plate can be observed. Between the observed locations, the route is reconstructed, assumed on certain goals, such as shortest or fastest route. An application of using license plates for route choice is given in Thomas & Tutert (2009). In all cases, only the chosen route is observed, which means that all other information deemed necessary for a specific route choice model needs to be generated or acquired in other ways.

Surveys as the other method for acquiring real route choices allows for a description of the chosen routes, but also enables researchers to acquire more data on, for example the choice set, the important attributes of route, motives for travelling and characteristics of the traveller. This method uses surveys (online, telephone based, paper based, etc.). As such, the data allows to provide a wider spectrum of information. Chosen routes in surveys however have a reliability problem, as it is difficult for participants to describe the route at the necessary level of detail. Often, certain parts of the route are left out, for which a choice have to be made how they are dealt with. Ramming (2001) uses data gathered using a paper survey and uses the shortest route to fill in the missing parts of the route, where the data was gathered using a survey. Another example of using an internet survey for acquiring route choice data is given in Prato et al. (2011), which uses an innovative way to acquire the data for the routes in an urban area, using a digital map with coded junctions. Also recently, a toolkit is developed which aids the researcher in the route choice analysis (Papinski & Scott, 2011).

Retrieving data on real route choice using surveys is a big burden for participants. It is a complex task to describe a complete route on paper or online, which has to be taken care of when designing a survey. Respondent fatigue can be an issue, causing unreliable and incomplete data. Combining the sources of data using a GPS tracking device and an online survey for additional data reduces the burden of the participants a great deal and seems to be the best way to acquire data on real route choices.

For SP data, it is obviously not possible to observe real life route choice behaviour. As a consequence, surveys or simulated environments are the only methods to get route choice data in the stated preference context. At one extreme there is the driving simulator such as those owned by TNO or the University of Leeds. These environments can be compared with a flight simulator and are often used to instruct new drivers or study driving behaviour in great detail, for example car following distance, merging, lane choice, etc. These environments also allow researchers to study new technologies such as adaptive cruise control in a realistic but controlled environment, whereas the same experiment in real traffic would introduce lots of issues, for example regarding safety and liability. A driving simulator provides an environment which is close to reality, as participants are actually driving a vehicle.

At the other extreme is the survey. Surveys do have different forms, but the most simple variant would ask respondents to select a route based on certain attributes and attribute levels. Such a survey is expected to have issues with validity, as it is difficult for participants to imagine the situation they have to make a route choice in. Next, the attributes used to describe the alternatives might be difficult to grasp for the participants, and they will never experience the consequences of their choices. Such a simple survey for route choice might not provide realistic results.

Figure 6.2: *Interface of the VLADIMIR route choice simulator (Bonsall et al., 1997)*

However, the internet and its dynamic webpages allow for interaction with the participant in a route choice experiment. It is possible to have them experience to some extent the consequences of their choices, particularly for travel time (by increasing the load time of the next webpage). It also allows for easy generation of the different scenarios in the experiment to get sufficient and equal amounts of data. Such a kind of survey is more of a controlled environment. An example of such an environment is the TSL (Travel Simulator Laboratory) as used in Chorus et al. (2007b); Chorus (2007); Bogers (2009).

A route choice simulator is an alternative that can be positioned between the full-scale driving simulators and the (interactive) surveys. Bonsall (2004a) defines route choice simulators as computer based tools for the collection of data on travellers' choice of routes. They involve asking individuals to make one or more trips through an artificial environment in which they are asked to make their way through the road network by choosing at each decision point. The route choice simulator keeps a detailed record of the choices. An example of the VLADIMIR route choice simulator (Bonsall et al., 1997) is given in figure 6.2. These simulators differ from the driving simulators, as the route choice simulators do not record the driving behaviour itself, but the route (and perhaps destination, modality and departure time) choices, while maintaining a certain degree of context as provided by a driving simulator. Compared to an (interactive) survey, the route choice simulator provides a more realistic view of the trip made, which make it easier for a participant to understand the situation and context of a trip. A route choice simulator combines the advantages of both SP and RP data, as it can observe "real" choices in a fictional situation which can be controlled by the researcher. Even though validity remains an issue with route choice simulators, they provide an environment which is closer to reality and therefor expected to be better understandable by participants.

6.2.3 Route choice simulators

In this research, a route choice simulator is chosen as the tool to acquire the data for the modelling of route choice behaviour during non-recurrent traffic situations in the presence of traffic information. This choice is based on:

- the need to include non-recurrent traffic situations such as incidents, road works or large events;
- the need to vary the traffic situations a driver experiences;
- the need to vary the type of traffic information a driver receives;
- the need to acquire the individual choices at each decision point (each junction) in the network, in order to estimate the model proposed in chapter 4;
- the need to use a cost and time efficient way to acquire the necessary data on route choice behaviour.

Route choice simulators are capable of simulating varying conditions, both in respect to the traffic situation and the trip purpose of the driver taking part in the experiment. The researcher can control the situations and make sure the participant will actually experience these conditions, especially in case of accidents. Using a route choice simulator it is also possible to create various kinds of traffic information and assure that a driver has received information on certain conditions. Using revealed preference data means that none of the above is possible. For example traffic information available in reality does not include information on accidents in urban environments, combined with information on congestion in urban areas.

Even though an SP approach using a survey would allow for a researcher to control the conditions and types of traffic information, it is a very difficult task to have a participant in such a survey experience the consequences of the choices or ask for a choice at each decision point. Route choice simulators are capable of tracking all choices at each individual location, together with the current traffic state of the network and traffic information available. This would also be possible using a full scale driving simulator, but these environments are costly. Next to that, experiments with driving simulators will take a long time if multiple trips are needed to be made with around 20 minutes of travel time.

A route choice simulator is thus the best approach to acquire data on route choice behaviour as it mimics reality but focuses on the choice data to be gathered in stead of driving behaviour data. Bonsall (2004a) concludes that route choice simulators have great potential as source of choice data. Compared to (online) surveys they offer:

- access to data on the choice process as well as on the choice itself;
- a means of allowing perception of key attributes to be part of the choice process rather than artificially excluded from it;
- a means to experience consequences of choices;
- reduced exposure to strategic response bias and unconstrained response bias;
- reduced exposure to good subject effects and to respondent fatigue or indifference effects, and;
- a mechanism for including a very wide range of potential behavioural stimuli.

Compared to a driving simulator, a route choice simulator offers:

- access to a wider range of subjects;
- considerably reduced costs, and;
- more rapid collection of data.

When looking at monitoring real world behaviour, a route choice simulator offers:

- much greater experimental control, and;
- easier access to a wide range of contextual data.

With so many advantages, route choice simulators have received a lot of attention in the study of route choice behaviour in the presence of information. An overview of capabilities of route choice simulators is presented in table 6.2, adapted and further completed with new route choice simulators from the overview made by Abdel-Aty & Abdalla (2006). The route choice simulator that will further be presented is based on the one presented in Raadsen & Muizelaar (2005).

The overview is based on characteristics and possibilities of route choice simulators. For this research, the important possibilities are the realism, en-route choices and multiple trials per subject. The realism can be further divided in a complex and realistic network and animation during the trips. These three necessary characteristics are further discussed in the design of the route choice simulator in the next section.

Table 6.2: *Overview of route choice simulators based on Abdel-Aty & Abdalla (2006)*

Authors	Year of publication	Network size	Realistic network	Animation	Different weather conditions	Internet based	Departure time	Mode choice	Pre-trip choice	En-route choice	Multiple trials per subject	Multiple users
Mahmassani & Stephan	88	S					X		X			X
Mahmassani & Herman	90	S					X		X			X
Bonsall & Parry	91	C					X		X	X	X	
Allen et al.	91	C	X						X	X	X	
Adler et al.	93	C	X						X	X	X	
Iida et al.	93	S							X		X	
Polak & Jones	93	NA						X			X	
Vaughn et al.	93	S		X					X		X	
Chen & Mahmassani	93	C		X			X		X	X	X	
Kantowitz et al.	93	C							X	X		
Yang et al.	93	S		X					X		X	
Koutsopoulos et al.	94	C	X						X	X	X	
Vaughn et al.	95	C		X					X	X	X	
Bonsall et al.	97	C							X	X		
Lotan	97	C							X	X		
Mahmassani & Liu	99	C	X				X		X	X		X
Chen & Mahmassani	99	C	X				X		X	X		X
Kraan et al.	00	C	X			X			X	X		
Ozbay et al.	01	C		X		X			X	X	X	
Mahmassani et al.	03	C	X			X			X	X		
Raadsen & Muizelaar	05	C	X	X		X				X	X	X
Abdel-Aty & Abdalla	06	C	X	X	X			X	X	X	X	
Avery et al.	07	C	X	X		X			X	X	X	
Bogers	09	C				X				X	X	

S = Simple network (2-node-2-link)
C = Complex network (more than 2 links)
X = Included

6.3 AIDA Route Choice Simulator

The starting point for the development of the route choice simulator for this research was the ability to incorporate accidents and their effects on traffic situations. It thus needs to be able to accurately model “normal” traffic, and represent this traffic to a participant. Another important element for the design of the route choice simulator was the ability to have multiple participants in the simulated environment at the same time. This would enable to acquire data on route choice behavior faster, but if enough participants would be in the same environment, it could also generate the effects described by Ben-Akiva et al. (1991):

- overreaction: too many users respond to the information, causing some roads to be busy and others to be quiet, continuously changing;
- concentration: too many users will use the same route, causing one route to be very congested.

However, for these effects to occur in a simulated environment a lot of participants are necessary. Such an amount would not work, without much extra care for the experimental setup. It is also not the specific goal of this research to investigate these effects. Next, the flexibility of the simulator was important. It should be easy to use different road networks, with different traffic situations, with for example accidents, road works or large events, without rebuilding the route choice simulator. The same applied to the type of traffic information and/or route guidance used in the simulator. These should be easily adapted to investigate effects of different kinds of route guidance or traffic information on the route choice. A combination of these requirements for a route choice simulator has not been found in any of the developed route choice simulator at the time of the design.

Traffic situations can be created by a range of models varying in temporal and spatial resolution, ranging from microscopic vehicle based models to macroscopic flow based models. Both are able to create realistic situations, calculate travel times, average speeds, and derive traffic information based on the traffic situation. In both types of models it is be possible to allow many participants simultaneously, although the workings will be different. Both types of models are also capable of simulating individual vehicles, where the macroscopic models use a dynamic approach and the microscopic models simulate individual vehicles. A microscopic model can be used for research into detailed behaviour as it includes individual movements of all vehicles. An advantage of a microsimulation model in a route choice simulator is the effect the controlled vehicles can have on the other vehicles in the road network. The other vehicles can adapt to the situation caused by the choice behaviour in the controlled vehicles, at any level in the simulation. A disadvantage of a microscopic simulation is the computing power needed to simulate an average size road network with lots of vehicles in faster than real time speeds. However, by providing a computer that is capable of running these situations, this can be circumvented.

Because of these capabilities and the level of detail, the choice was made to base the route choice simulator on a microscopic traffic simulation model. Such a model allows to create the various traffic situations, while using the abilities of the simulation model for moving the vehicles through the network, gathering the necessary data for traffic information and storing the traffic data for further analysis. At the time of the choice for a simulation model, Quadstone Paramics was one of the few that offered an interface (Application Programming Interface or API), with the most capabilities to access the inner workings of the microscopic simulation model.

For the design of the route choice simulator, further on called the AIDA RCS, much attention has been given to the work done in Koutsopoulos et al. (1995); Firmin (1995); Bonsall et al. (1997); Bonsall (2004a); Chorus (2007); Bogers (2009). All these studies investigated the use of route or travel simulators for acquiring data, which was otherwise impossible or very difficult to retrieve. Especially the overview presented in Bonsall (2004a) raises the issues that need to be addressed when designing a route choice simulator or an experiment based on a route choice simulator. These five issues are:

1. whether to represent a real network;
2. how to set traffic conditions in the simulator;
3. how realistic should the interface be;
4. the inclusion of questions to supplement the decision log, and;
5. how important is portability.

The items 2 to 5 are addressed in the design of the AIDA RCS, where the first item is part of the design of the experiment, not of the simulator. Both types of networks (a real, or a not existing) can be created in Paramics. The traffic conditions in the AIDA RCS are set by using a microscopic traffic simulation model. This model allows to vary the traffic conditions between several trips, either by using varying random seeds or completely different OD matrices. And because many participants can simultaneously make a trip in the simulator, it is even possible to have the traffic conditions reflect the choices of these participants. The choice made regarding this item can be found in section 7.2.1.

Item three about the realism of the interface concerns not just the drivers view, but also the presence and/or complexity of the driving task and the representation of time in the simulator. A view for the driver which is "natural", or closely resembles the view from an actual car, is found to be important (Firmin, 1995). However, the extent to which extra information is included, such as a preview of upcoming links, influences the choices by the participants to a certain extent. For the AIDA RCS it was chosen to include a view of the dashboard and windscreen, and present a map of the road network, including the traffic conditions. The interface should also be able to present traffic information and a route advise in a realistic way.

Based on the results of the study of Firmin (1995) it was chosen not to include an alternative driving task in the AIDA RCS. In a route choice simulator, a participant does not have to steer, accelerate or brake. This means that the participant would have more time for making a choice. An alternative task, such as keeping a square in a certain area of the screen with the keyboard, could represent a driving task. However, such a task might become more important than the task of choosing a route. Therefore, the AIDA RCS does not have such a task, but uses the possibility to speed up the time.

The third element of realism is the representation of time. Previous research (Bonsall et al., 1997; Bonsall, 2004a; Bonsall & Palmer, 2004; Firmin, 1995) indicates that a link between elapsed time and time in the simulation is important. If a simulator has no such link, it might be very well possible that a participant is not making choices to minimize the time in the simulation, but the elapsed time, which is unwanted behaviour. However, if the simulated and elapsed time are equal, experiments will take too much time and thus mean a large burden on the participant, particularly because no alternative driving task is to be performed. The AIDA RCS therefore aims at speeding up the simulated time to a certain amount. Simulation time is thus proportional to the elapsed time by a constant:

$$\text{Simulation time } = c * \text{ Elapsed time} \tag{6.1}$$

in which c is a constant between 0 and 1.

Related to the realism of the simulator is the realism of the experiment. This is important for the validity of the simulator and the experiments. An important element in this validity is the inclusion of costs. For example, if a participant selects a long route each time, the experiment takes more time. A participant however only experiences this longer duration, but does not experience the effects of a slow route as he would have done in reality (too late arrivals at work or meetings, missed events). Such an undesired effect can be prevented by including costs. For example, for every minute a participant would be too late at the designated location (in a commute scenario it would be a work location) a certain amount will be deducted from the participant's fee. However, what amount would be deducted? Which weight to impose? These values are difficult to obtain and it is not agreed upon that the inclusion of costs in the experiment yields better results (van der Mede & van Berkum, 1991). Contrary to other experiments with route or travel simulators, it was chosen not to incorporate such costs in the AIDA RCS. It is expected that participants are able to make realistic choices if the scenarios in the experiments are properly set up and if the simulator itself makes the trip realistic. This last assumption is targeted by the design of the AIDA RCS, while the scenario's are part of the experimental design (see chapter 7).

The inclusion of questions to supplement the decision log is also addressed in the AIDA RCS. Based on the conclusions made by Bonsall (2004a), it is important to ask questions to participants of a route choice simulator study, because such a simulator only allows for passive observation. It thus does not observe the reasoning behind the choices made. A number of questions that do not interrupt the simulation itself allow for more or better of the observations. The inclusion of questions and their answers with the observations from the route choice simulator are thus very important. Possible questions are about the validity of the route choice simulator, by asking participants about their experienced realism and if they would make an equal choice in reality. Of course, the answers, as well as the observations should be stored and be easily available for analysis.

The portability of a route choice simulator is also important, because it makes it easier to attract different participants to the study. A full-scale driving simulator for example is usually located at a car manufacturer, university or a research institute, which makes it more difficult to attract a range of different participants.

A portable route choice simulator can also be used at different locations, for example office or factory workers. This ensures a higher representation of the different groups in society. It does come at a cost however, because certain items such as multiple screens and sound are more difficult to be portable. Another issue is that the location cannot be as controlled as in a dedicated research room. For our research purpose, portability was considered to be more important than a controlled experimental environment. The AIDA RCS is therefore chosen to be a portable simulator. Another advantage of the AIDA RCS would be that it is possible to participate without actually being present at a certain location. Because of the advanced networking capabilities of today's computers, it is possible to participate over the internet, using a client/server architecture.

6.3.1 AIDA RCS architecture

Based on the issues described in section 6.3, an architecture for the AIDA RCS was made (see figure 6.3). The architecture shows the different components that are used in the AIDA RCS. The initial design was done by Raadsen (2005) and further developed for this research.

On the left side the Paramics Modeller (Quadstone, 2007) is placed. The Paramics Modeller is part of the Paramics Suite (of which version 5.2.2 was used) which consists of several parts:

- Modeller: tool to create the road network and run the simulation (with graphical user interface);
- Analyser: tool to analyse data generated by the simulation;
- Processor: runs simulation without graphical interface to speed up the application;
- Programmer: contains the developer API;
- Monitor: interface to monitor the generated pollution;
- Estimator: contains the OD estimation model;
- Designer: the 3D design tool;
- Viewer: a viewer for the created network.

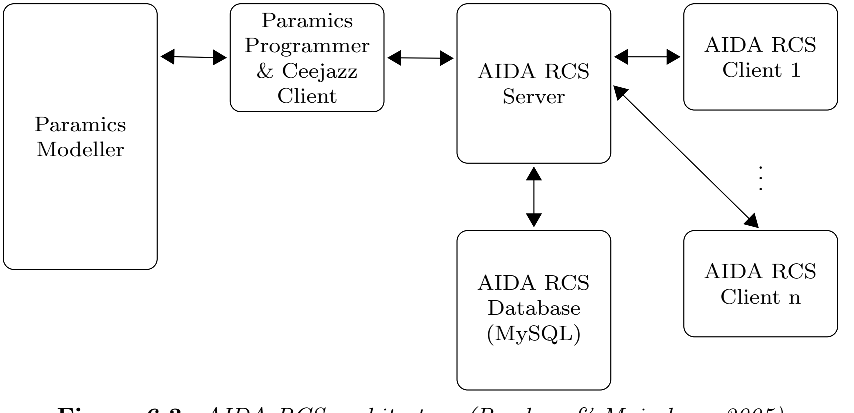


Figure 6.3: *AIDA RCS architecture (Raadsen & Muizelaar, 2005)*

The modules of interest are the Modeller module to design the network and to run the simulation, and the Programmer module which is used to connect the plug-in to the RCS. The programmer is the connection to the outside world as it provides functionality to write plug-ins. This API is split into four functional parts:

- QPO: Override function, if these functions are implemented they will override the default Paramics procedures. Examples of these kind of functions are:
  - car following
  - lane changing
  - route choice
- QPX: Extended functions to add extra functionality to the current behavior like:
  - start/stop simulation
  - detection of released vehicles
- QPS: These are the "set" functions that can be called by the plug-in such as:
  - set the speed limit on a link
  - set the current speed of a vehicle
- QPG: These functions define the provide "get" functionality to retrieve important data from the simulation like:
  - the number of vehicles on a link
  - if a car is involved in an incident
  - what the destination zone of a vehicle is

The API available from Paramics Programmer requires making plugins in C++, as .dll files. There is also an interface to this API available for the Java programming language, which is the Ceejazz Client (Azalient, 2007). For the AIDA RCS, the Ceejazz client is used, because of the advantages of Java.

In addition to the use of the programmer API, there is another source of manipulating a Paramics network. Paramics uses ASCII files to store all data defining its road network and traffic situation. These files can be interpreted and manipulated by the AIDA RCS to extract data on route guidance, special events, incidents etc. without being limited by the Paramics plug-in functionality. Each Paramics network consists of a number of files. These files can be categorized into five different groups; each group has a number of files. The files which are of interest to the AIDA RCS are indicated below:

- Network
  - Nodes
  - Links
  - Categories
  - Priorities
- Demand
  - Demands
  - Incidents
  - Profile
- Infrastructure
- Statistics
- Initialisation

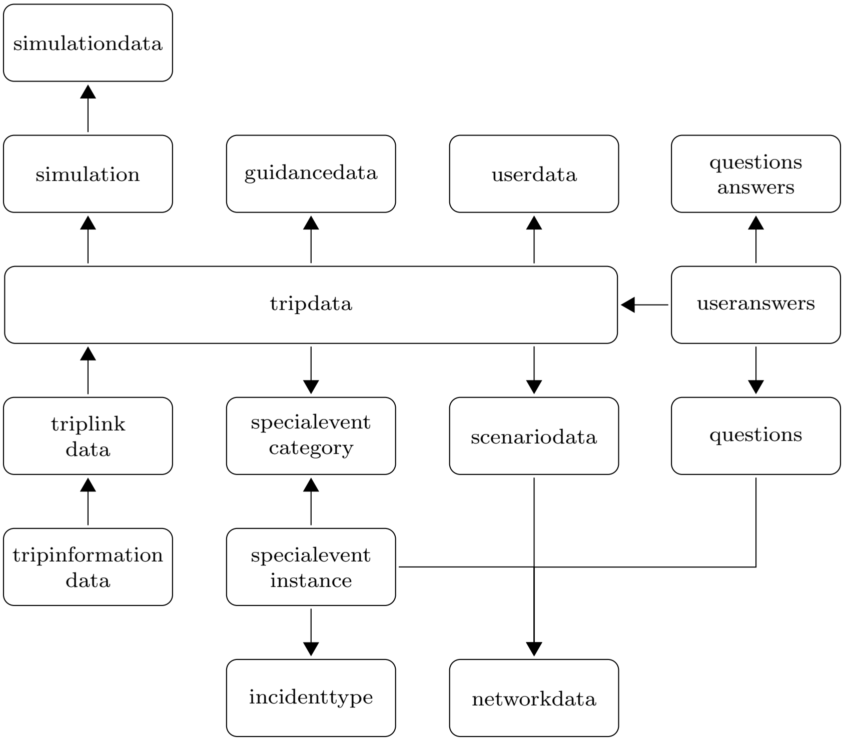


Figure 6.4: *AIDA RCS Database structure*

Based on the requirements for storing data on trips from the simulator, as well as storing data from questions before and after trips, Paramics itself could not fulfill the storage needs. The easiest solution for storing data of the experiments is using a database. The requirements were translated into a database design (see figure 6.4). This database design shows several tables related to the simulation (storing data of the trip, the links and the information), several tables related to before the trip and/or the simulation (storing information on the user, and selecting a scenario and special event) and tables related to after the trip (storing data on the questions). In other words, the database is also used to configure and set up the scenario's.

For the server and clients, the Java programming was selected, although also C++ was available. The main advantage of Java over C++ was the availability of communication between the server and the clients. For the database, MySQL was chosen. MySQL is freely available and supports concurrent access, which is important in the design of the server, because the server needs to store data and retrieve data simultaneously. Communication between Java and MySQL is possible by the MySQL JDBC driver.

(a) *No message* **(b)** *Traffic information*

Figure 6.5: *Driver's view of the AIDA RCS*

6.3.2 AIDA RCS Client

The client is the interface of the experiment with the participant, as described in section 6.2. The graphical user interface or driver's view of the AIDA RCS is represented in figure 6.5. In these screenshots of the simulator, three distinctive elements can be found:

1. the windscreen with a view of the upcoming junction;
2. the dashboard with speedometer, and;
3. the navigation information.

The windscreen shows a picture of the upcoming junction. It is possible to use realistic pictures or screenshots from Paramics (in its 3D view). It loads the background directly when a vehicle starts on a new link, and is static. It is also possible to show a picture even if the junction only has one exit (straight ahead), because it is based on the network files of Paramics. In other words, for each node and each link leading to that node, a picture can be shown to the participant. If no picture is available, the windscreen remains black.

In the bottom of the windscreen, the congestion on the current link is shown, based on the density on that link and converted into three categories (see table 6.3). These categories are also related to three colors (used throughout the client): (i) green for freeflow; (ii) orange for medium traffic; (iii) red for heavy traffic. The windscreen is also used to display messages (figure 6.5b). The contents of these messages are either on traffic jams or special events, such as accidents.

The dashboard is always shown once the participant starts the trip. It shows the current speed of the vehicle that is selected, which is controlled by the car following model of Paramics. On the dashboard the current time in the simulation is also shown, taken directly from Paramics. Based on the type of traffic information that is selected for the trip, the estimated time of arrival is also shown, below the current time. This estimated time of arrival is based on the current fastest route with the current travel times of each link.

The large area on the right of the screens is the information that would normally be presented in the personal navigation device. For most car drivers, the place of such a device is to the right of them (except perhaps for countries which drive on the left), hence the navigation information is presented on the same location in the AIDA RCS. The area consists of four elements:

1. a map of the road network;
2. a map of the upcoming junction;
3. a list of the roads, and;
4. a blue bar indicating the time left to choose.

The map displays the origin and destination, both as blue dots. The current location of the vehicle is shown as a white dot. The colors of the links are again based on the density of these links, but there is an issue with these colors. Because most roads are two-ways streets, there are two links between nodes (apart from on-ramps). However, the map is to small to display these two opposing links next to each other. Only one of these links is shown on the map; the link that is drawn last. Which link is drawn last and thus overlaps the opposing link is based on the way the road network is constructed and read into the RCS. The order in which the links are parsed (and thus are drawn) is based on the start node of the link. Because the way in which the nodes are constructed and numbered cannot be controlled, eventually, the order in which the links are drawn cannot be predicted or controlled. This implicates that although the density of the links is shown in the map, it is not clear which direction is actually shown. It gives an indication of the congestion status of a link, but a participant cannot be certain about its direction. The map does not show any route guidance if that information is selected. The map is also not rotated towards the orientation of the driver.

The map of an upcoming junction is rotated to the orientation of the driver, and always shows the approaching link on the bottom. This resembles the view from the windscreen and the normal view in personal navigation devices. The exits to choose from are proportionally drawn, based on the angles from the map of the complete road network. The exits are numbered, and the exit that is currently selected by the participant is made blue. If route guidance is selected, the advised exit is also drawn with an arrow at the end, indicating the best option at that moment. The exit links are also given a color, again based on the three congestion levels defined in table 6.3.

The list of roads show the names of the exits to choose from, with the same numbering as the map from the upcoming junction. The selected exit is again shown in blue. In front of the exit numbers, three small triangles can be seen.

Table 6.3: *Congestion categories, based on (May, 1990, p. 194)*

Density (vehicles/km)	Level of service	Flow conditions
0-8	A	Uncongested or freeflow
8-12	B	
12-19	C	
19-26	D	
26-47	E	Near capacity or medium traffic
47-62	F	Congested or heavy traffic
>62		

These indicate again the congestion level on that exit, in the form of a color and an extra triangle. On top of the exit names, the current link is always shown. Depending on the selected type of traffic information, also the advised link is shown, together with its number, which allows for easy selection in case of multiple roads with the same name.

A blue bar is displayed below the list of roads. This blue bar indicates the time left to choose, before the currently selected exit is communicated to the server and to Paramics. Because a certain time is necessary to parse the selected exit and communicate it to the server and change the route choice for the vehicle in Paramics, the time is less than the total time a vehicle needs to traverse the link. This is also due to the fact that the client does not continuously have information on its current position, is this information is only communicated at regular intervals. Hence, the threshold distance that is used in the client is based on equation (6.2), in which the buffer time needs to be larger than the regular update interval (Raadsen, 2005).

$$\text{threshold distance} = \text{speed limit} \times \text{time ratio} \times \text{buffer time} \tag{6.2}$$

$$\text{time ratio} = \frac{\text{number of steps}}{\text{real time difference}} \tag{6.3}$$

in which:

threshold distance (m) = the distance at which the selected exit is communicated
speed limit (m/s) = the speed limit of the current link
buffer time (s) = the amount of time the client has before it has to communicate the choice
number of steps (s) = the amount of simulation steps taken (equal to the simulation time step, which is 1 second)
real time difference (s) = the time difference in reality between the simulation steps

If a participant does not make a choice in time (the blue bar reached its end), it is not possible to stop the simulation and force the participant to choose. The fact that multiple clients can be connected, makes it impossible to stop, because all participants would see no progress or updates in their clients, while they do not see any reason for that behavior. Thus, it is necessary to make a choice in order to allow the controlled vehicle to continue its path. This choice is Paramics choice; it is based on the link that Paramics would have chosen for this vehicle based on its own routing algorithm. This routing algorithm is further discussed in section 7.2.1. However, it is recorded that this choice was not a user choice, to able to distinguish real from automatic choices in the analysis.

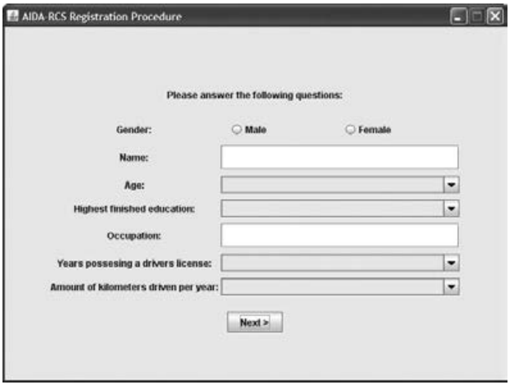


Figure 6.6: *AIDA RCS Client pre trip screen for entering personal data*

As already mentioned, navigation is part of the client. Navigation takes care of calculating a specific type of route from origin or current location towards the destination. The client is capable of using multiple navigation algorithms to decide on these routes, by allowing a configuration to select the specific type. In the AIDA RCS currently only one routing algorithm is implemented, which is the A^* algorithm (Hart et al., 1968). In the A^* algorithm, the heuristic costs between nodes is based on the crow fly distance between nodes and a speed of 50 km/h. For the actual costs between two nodes, the current travel time is used, which is the freeflow travel time with the delay time measured in Paramics.

The results from the A^* algorithm are twofold. First, the results are used to inform the participant if the scenario involves dynamic route guidance. Each new link, the route is recalculated, as well as when the traffic situation from Paramics is updated. The user is only informed about the advised next link, based on the current fastest route. Second, the A^* algorithm is also used to calculate the fastest route towards the destination for each exit from the current link. In other words, all options on a junction are evaluated for their fastest route. These routes, including the overlap between them, their specific total and freeflow costs and their lengths, are sent to the AIDA RCS Server for storage in the database. This data can be used for analysis of the route choice.

As part of the complete sequence in the AIDA RCS client, each trip starts with pre-trip questions. An example of such a screen is found in figure 6.6. All screens can be found in appendix C.

The pre-trip questions asked to the participant concern:

- gender;
- age group;
- level of education;
- current occupation;
- years owning a drivers license, and;
- amount of kilometers driving per year.

These questions are only asked once, when a new participants registers in the AIDA RCS. Subsequent trips can be made with just their name and user ID. Once a participant is registered, a choice has to be made regarding the scenario and the type of traffic information. Finally, the selections are returned to the participant in order to ensure the right choices.

Once the participant has arrived at any destination (whether it is the correct destination or not, more on this in section 6.3.3), he is informed about his departure and arrival time, and the travel time, followed by a set of questions. These questions and their answers are easily changeable because they are communicated from the AIDA RCS Server. The questions are available in the database, and once a car has arrived, the server retrieves the proper questions and answers from the database and sends it to the client. The chosen answers are then send to the server, which stores it in the database.

6.3.3 AIDA RCS Server

The AIDA RCS Server takes care of the communication between the database, the clients and Paramics. The server is started from within Paramics, because this is the easiest way a plugin (which the AIDA RCS Server is for the API) can be run. The server needs to be configured on four elements (figure 6.7):

1. network (defines the update intervals and buffer time, as well as information for the clients to be able to connect);
2. simulation (describing the setup of the current simulation, including minimum lengths for traffic jam detection and communication);
3. database (describing the setup of the communication with the database), and;
4. special events (defining if and which special event will take place during the simulation).

Communication between the server and the clients takes place on two instances. The first instance is the regular data. Regular data consists of information on:

- current vehicle speed (m/s);
- distance left until end of the link (m);
- current ratio between elapsed and simulation time;

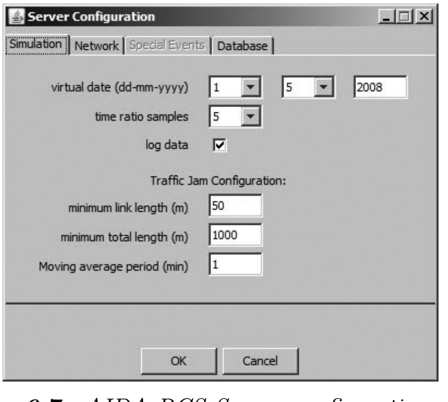


Figure 6.7: *AIDA RCS Server configuration screen*

- current geographical position of the vehicle (x, y and z coordinates);
- density of the current link (vehicles/km);
- destination reached boolean, and;
- special event related messages.

The second instance is a semi-regular data communication. This data concerns:

- number of vehicles on every link of the network;
- delay in seconds of every link of the network, and;
- all traffic situation related messages at the current time;

This separation of data packages that are communicated is based on the size of the second instance. The semi-regular data is of much larger size than the regular data. The interval of both instance needs to be defined before the simulation and cannot be changed meanwhile. It is important to choose the intervals carefully, because for example, the semi-regular data informs the participant of congestion. If the interval is very small, the participant can get traffic jam messages every few seconds. It is also dependent on the size of the network and the rate of change for congestion on the various links. Another issue that needs to be taken into account is the latency of the network between the client and the server. For the regular update interval, the buffer time is also important, as described in equation (6.2). The default for the regular package is every 2 seconds, while the semi-regular package is send every 60 seconds.

Both the semi-regular and regular data package have an element which needs further explanation. The regular data package contains information on special events, which are configured before starting the simulation. The original setup enabled three types of special events: (i) large events; (ii) road works, and (iii) incidents. For a large event, the OD matrix has to be adjusted. Road works means that a link will be turned into a one-way street, or completely closed. For incidents, Paramics has the possibility to create an incidents file which describes the time, duration, location and effect of incidents. However, these files can only be loaded when the network is loaded in Paramics. This means that it is not possible to dynamically create incidents, but only use a preselected incident. This incident needs to be described in the database (see tables concerning special events in figure 6.4).

An important element in the semi-regular data package concerns the traffic situation related messages. These are basically messages on traffic jams in the network. The traffic jams are detected on a link by link basis in the AIDA RCS Server when the density is higher than 47 vehicles per kilometer (see table 6.3). Paramics in itself is not capable of defining traffic jams, which meant that the detection and description of traffic jams needed to be realized in the AIDA RCS. For this purpose an algorithm was made, which checks for links larger than a certain distance, and if this link has a density larger than mentioned. These links are entered in a list. Next every link is checked for exit links which also have a higher density. These links then are joined together. However, this also presents problems for links which have two or more links upstream that are also jammed. A choice needs to be made to select which of these links belongs to the same traffic jam, whereas the other is a new traffic jam. This rather arbitrary selection is necessary to prevent double counting of the same link as a traffic jam. The same issue applies to downstream jammed links. Currently, for a T junction, the exit with index one is chosen. For a four legged junction, the exit with index two is chosen, which is usually straight ahead. Neither choice takes the actual road network layout into account, and is independent of type of road.

Once a full list of traffic jams is created, these traffic jams are checked against a defined length, which is configured when the AIDA RCS is started. Traffic jams shorter than this length are discarded. The other traffic jams are communicated to the clients, in the semi-regular data package. These traffic jams are also stored in the database, whereas the shorter traffic jams are not stored. It is not possible to see where short traffic jams have occurred during the simulation, and if these short jams could have been connected to another longer traffic jam.

The traffic jam algorithm is not necessarily correct or complete. For example, there is no check on a gridlock situation, where a traffic jam forms a complete circle. Such a situation would now lead to an infinite traffic jam, which obviously is not right. As already mentioned, a not necessarily correct choice is made for the connection of links with traffic jams. This could lead to a situation where the traffic jam connects a highway link to an off-ramp link, while the most obvious choice would be to connect the highway link to another highway link (when both are jammed). This needs to be taken into account when analyzing the influence of traffic jam messages. However, it is not possible using the stored data to check if these kind situations have occurred.

Paramics, as a microscopic simulation model, simulates individual vehicles. At every time step in the simulation, for every vehicle certain parameters are calculated, such as its speed, lane choice, and next link. Since the AIDA RCS is targeted at route choice at link level, it is necessary to control the next link of vehicles. To this end, when a client connects to the server, a vehicle needs to be selected, and brought under control of the AIDA RCS while Paramics will maintain the control over the speed, lane, etc. of the vehicle. The participant now controls the next link of this selected vehicle. This presents two problems. First of all, a vehicle cannot be created dynamically in Paramics. This means that it is necessary to wait for the normal release of a vehicle, which is based on the OD matrix in Paramics. If very few vehicles depart from that location, it can take quite a while before a vehicle can be taken over. Every vehicle that departs is checked for its origin and its type. If the origin and type match (only normal cars are selected), a vehicle is taken over. Only when a vehicle is taken over, the trip in the client actually starts. A participant will be able to notice such a delay. The other problem in Paramics is that a vehicle does not necessarily start at the end of the link. This is due to the way Paramics releases vehicles from an origin. Origins in Paramics are defined as zones, which can have more than one link. This means that Paramics can arbitrarily choose a link to which to release a vehicle, but also on which exact location on that link, as long as it inside the zone. When one link exists inside the zone, Paramics randomly distributes its starting location on that link. This is not what a participant would expect. Therefore, once a vehicle is released and claimed, it is immediately moved to the starting point of the link. This is taken care off by the AIDA RCS Server.

Every time a choice is made, this is parsed by the AIDA RCS Server, communicated with Paramics and stored in the database. Normally, a vehicle in Paramics makes his choice for a link, two links ahead. This allows the vehicle to choose his lanes on these links appropriately. However, for a route choice simulator, this means it is difficult to route a vehicle through the road network properly. Luckily, the API allows for a last moment call of the function that determines the next link. It means that lane choice might not be the best possible, but a participant usually does not see the consequences. In some cases, it might result in a longer delay before the controlled vehicle will move to the chosen link, because a queue exists. Roundabouts in Paramics are however more difficult, because a roundabout is made of several sub links. Each of these links needs to be chosen first, before the final exit can be chosen. This is quite impossible for a participant to do, and very unrealistic. However, in the AIDA RCS it turned out to be very difficult to choose the appropriate route over the roundabout. Because of this difficulty, roundabouts are left out in the AIDA RCS.

Once a claimed vehicle has arrived at its destination, a flag will be set. The client controlling that vehicle will notice this destination-reached-flag and start the post-trip phase. However, it is not guaranteed that a participant will arrive at the proper destination. It is possible that a dead end street is chosen, different from the actual destination. Paramics will detect that this vehicle reached a zone, and remove the vehicle from the network. No functionality is available to instantaneously put the vehicle back on the link in the opposite direction. This means that even when the wrong destination is reached, the post-trip phase will commence for the client. At the end of the trip, the data of the trip and the post-trip questions are stored in the database.

6.3.4 ATIS

The AIDA RCS is developed to collect data on route choice behaviour under information provision. Because it is a virtual environment, it is possible to define new types of information. Paramics as the component in the AIDA RCS architecture providing the road network and traffic situation is also capable of providing numerous details of the network, the traffic situation and individual vehicles. All such data can be used to create traffic information.

For the AIDA RCS two elements are previously described that are of importance for traffic information. The first is the semi-regular message between the server and the clients. This message contains all traffic situation related message for a current time stamp. In the used implementation of the AIDA RCS, these message can be used to communicate traffic jams and information on accidents, but also other events can be used. This message contains the links it concerns, a possible text describing the event and in case of traffic jams the total delay, average speed and length. The last element is the status of the event, whether it is new, already reported, or cleared. This data is stored in the database, while the client processes this data to create a message for the participant in the windscreen. An example of a displayed message is: “Traffic jam on Ring; standstill; NEW”. These traffic messages, either on events or traffic jams, are defined as Textual information.

Another type of traffic information, extended the Textual information. It uses the same information and presents this information in an equal manner. It is extended with navigation. The other elements in the semi-regular message contain information about the average speed and delay per link. This information is used to provide a route towards the destination based on the smallest current travel time. To this end, the A* is used to find the fastest route in a short amount of time. This type of traffic information is defined as Dynamic Route Guidance.

6.4 Validity

The AIDA RCS was developed using the requirements and the issues mentioned for the design of a simulated environment in Bonsall (2004a). These issues are important as they are likely to influence the validity of the resulting data. The validity of these environments has been discussed in Koutsopoulos et al. (1995); Firmin (1995); Bonsall et al. (1997); Palmer et al. (1998); Mahmassani & Jou (2000); Bonsall (2004a); Chorus et al. (2007b); Bogers (2009).

Koutsopoulos et al. use the following definition: “Validity of a simulator refers to the correspondence between the results acquired by using the simulator and a set of outcomes that are needed or desired and constitute the objectives of its use.” This resembles external validity and relates to behaviour of drivers during non-recurrent traffic situations while they are presented with different forms of ATIS, compared to the same behaviour in real life. Such comparison is only possible if real life data for the same situations is available. As this research is focussed on non-recurrent traffic situations, especially in urban environments with traffic information on these situations, the real life data does not exist. This makes it impossible to verify the results of the AIDA RCS as a form of external validity.

However, as mentioned by Mahmassani & Herman (1990) external validity concerns the relation between the results which are obtained by experiment and the real world, or in other words the transferability of the obtained results. They divide the results into four elements which can be transferred:

1. principal theoretical constructs, which concerns theoretical developments which are necessary before data can be obtained;
2. methodology for model specification and estimation, which concern the formulation of models and their estimation that capture the appropriate behaviour dynamics;
3. behavioural insights, which concerns the relative importance of various attributes or variables in an estimated model, and;
4. specific models, which concerns the application of the same model and methods in a real life and simulated environment where the relative magnitudes and signs of parameters are of interest, opposed to the individual values.

Validity can be separated to three types:

- external validity;
- internal validity, and;
- face validity.

External validity is a a measure for generalization of the (causal) inferences based on the experimental data to real life. Internal validity defines the degree of causal inference that can be made. High internal validity means that there is a (statistically) proven relation between a cause and effect for which no alternative plausible explanations are available. Face validity describes wether the experiment looks valid and is quite often based on self-reported assessments by participants and intuitive judgment by the participant or research. Face validity is the weakest form of validity, but together with internal validity often the only form available if no real life data can be used to compare experimental findings.

All types of validity mentioned have been studied for other route choice simulators. Internal validity is often the result of the model development itself in terms of test for alternative hypotheses. Bonsall et al. (1997) tested the validity of their VLADIMIR route choice simulator on one aspect face validity and two aspects of external validity. Face validity was based on questions asked to participants after completing their journey. They were asked if they think they would have followed the same route if they had been making the same trip in real life, together with possible reasons for discrepancy between the trips in VLADIMIR and their real life trips. The network used in VLADIMIR was a real one (North-West Leeds), for which participants were provided a realistically specified trip. Overall, the trips made without guidance are given high ratings (between 70% and 100%) for making the same journey in real life whereas trips with guidance show a lower level (around 50%). Familiarity with the road network was found to be important in the confidence in the chosen routes in the simulator. As such, guidance might cause participants to experience less familiar areas which means a lower confidence in their route choice.

The external validity is investigated by comparing models based on simulator and real life data, as well as the test whether choices in the simulated environment are indistinguishable from choices made in equivalent real life situations. The results for the first part show that there is hardly any difference between the models from both sources of data. Combining the models with a different scale parameter also shows there is no significant difference in variance of the underlying error. Leaving out trips where a participant stated a low confidence in making the same route choice in reality, does not have a big impact on the models or the underlying variance of the error. The data from a route choice simulator thus are very well applicable to real life route choices. The comparison of routes made in real life circumstances with route made in VLADIMIR with equal circumstances neither show a difference for participants with a high degree of familiarity. For 54 journeys made, 47 followed the exact same route in the simulated environment. This addresses the third and fourth element of the list by Mahmassani & Herman (1990) and indicate that both specific models and behavioural insights based on VLADIMIR as a route choice simulator are very well applicable in real life circumstances. This applies to both guided and unguided trips.

Mahmassani & Jou (2000) investigate all the elements regarding the transferability of findings. They conclude that a behavioural framework which was developed for day-to-day commuter behaviour was valid for both experimental and real life observations, where the resulting estimated variables are having valid signs for different observations from both real life and experimental conditions. The parameter values themselves are not identical to real life situations, in which it differs from the findings of Bonsall et al.. However, the variance is of the same range, which indicate the model can be compared for both data sources.

Chorus et al. (2007b) focus on face validity and convergent validity as a specific form of external validity. In this research a multimodal travel simulator with next-generation ATIS is used to study travel decision making. In this case a purely fictional network is used. The validity was first studied using a self-reported judgment by the participants. They indicate that the environment was easy to understand and they had no difficulty identifying with the different (fictional) travel situations. Next, the face validity and convergent validity are determined using intuitions. For face validity they look at the kind of choice that would be made in real life under comparable conditions, where convergent validity is studied by investigating if the choices relate in meaningful ways to choices made in real life under different conditions. The latter suggests that some elements of use and effects among travellers for currently available ATIS may relate to the next-generation ATIS. As it concerns intuitions, they were aware of the fact that new insights should not be ruled out because new technologies might show use or effects that are counterintuitive.

The convergent validity is based on the concept that car users in real life are more likely to choose the car option in the simulated environment. It does not directly relate the choices in the simulated environment to those in real life, but is based on consistent behaviour. Next, travellers who acquire more information in real life are also more likely to acquire more information in the simulator. The face validity is based on choices for modality, information acquisition and information effect. They relate for example the travel time of a mode to the probability of being chosen. In both cases, the intuitions are shown by comparing data based on current behaviour (acquired as revealed preference data in the survey). Both types of validity are concluded to be high, as the intuitions formulated are statistically proven. For example, car users were indeed more likely to chose the car option. Chorus et al. (2007b) conclude that although the environment concerns hypothetical situations and not yet existing types of information, the environment appears to generate valid data.

Results from previous research using route choice simulators suggest that a well designed choice simulator will result in data and observations that are valid. These studies also show that varying model frameworks can be developed, estimated and shown to have valid results, either with similar parameter values or with relative values and appropriate signs. They have shown to produce valid results for various types of traffic information and guidance, and even show a high level of validity for results in multimodal environments.

6.5 Summary

For this research, stated preference data is selected as the best alternative, as it is possible to combine the necessary context of route choices and control this context. The main issue for such data is the validity. Data generated using predetermined models is not possible in this research, as the goal is to create a model that provides behavioural insights for route choice. Generated data are better applicable to show certain new approaches for route choice or model estimation. Revealed preference is neither useable for this research, as the combination of traffic information and non-recurrent traffic situations is impossible to control in real life, where it would also be difficult to acquire sufficient data on the context of the observed choices.

Simulated environments have been used for the generation of stated preference data regarding traveller behaviour. These environments have received much attention in the field of ATIS and route choice behaviour, but also have seen applications for multimodal choices. An overview of the capabilities of many of these environments is given in table 6.2. Using the recommendations and issues in Koutsopoulos et al. (1995); Bonsall (2004a) the AIDA RCS was developed. The AIDA RCS is based on a client-server architecture, where the server connects to both Paramics for its microscopic traffic simulation and road network and a relational database for storage and logging. The server allows to be connected over any ip network, making it capable to be used over the Internet. Multiple clients can be connected simultaneously and will be travelling in the same traffic simulation. The client itself has a windscreen and dashboard, and shows various information regarding the location in the network, the upcoming junction and time left to make a choice.

The AIDA RCS allows for pre trip and post trip questions in which it is possible to ask participants for their real life behaviour as well as their opinion about the environment and trips made herein. It is capable of representing real networks and realistic traffic situations, although it is possible to create any kind of road network and traffic situations, using Paramics to create these. The interface is not as realistic as it would be in a driving simulator, but resembles the location of information as often found for vehicles in a right-hand driving country. In the AIDA RCS an link between the time travelled and the time elapsed in reality is maintained, with the aid of a speedup factor to prevent overly long trips. The link increases the realism as participants will experience the consequences of their choices with longer travel times and longer elapsed times. Because of this link and possible issues, no incentive structure in the AIDA RCS is used. Various other studies have shown that simulated environments and route choice simulators in particular have a high degree of validity, in terms of internal, external and face validity. This lead to a list of five issues which should be addressed when designing a route choice simulator. Because the AIDA RCS has been designed with this list in mind, including the importance of valid results, it is expected that a high level of validity is achieved for this research.

The AIDA RCS is used in an experiment to determine the impacts of both non-recurrent traffic situations and traffic information on route choice behaviour. The experiment design is further discussed in chapter 7. The results of the experiment are presented in chapters 8 and 9.

What's in store for me in the direction I don't take?

— Jack Kerouac

Chapter 7

Route choice experiment - Setup

7.1 Introduction

In this chapter we describe the experimental set-up of the route choice simulation experiments. We focus on route choice behaviour during non-recurrent traffic situations caused by accidents. We use the AIDA Route Choice Simulator described in chapter 6. The setup of the experiment and its elements are described in section 7.2. The results from the experiment are presented in chapters 8 and 9. The chapter will finish with a summary in section 7.3.

7.2 Experiment setup

Based on the capabilities and design of the AIDA RCS a route choice experiment was setup based on a number of requirements. Recall the five issues given by Bonsall (2004b):

1. whether to represent a real network;
2. how to set traffic conditions in the simulator;
3. how realistic should the interface be;
4. the inclusion of questions to supplement the decision log, and;
5. how important is portability.

The items 3 to 5 have been addressed in the design of the AIDA RCS (see section 6.3). The first two of these issues are elements that will be discussed here.

The AIDA RCS is a virtual environment based on Paramics capable of simulating a wide range of different road networks and traffic situations. To focus the development of both these elements in the environment, a high level set of requirements are defined, for which the guidelines from (Bonsall, 2004b) regarding the network and traffic conditions, have been used. These are combined with the scope of this research. The requirements are as follows:

- the road network should consist of varying road categories, at least including highway, provincial, and urban road with matching speeds and lanes;
- the road network should have a city centre and suburbs;
- the road network should use no roundabouts or traffic lights, as these cannot be used in the AIDA RCS;

- the traffic situation without accidents needs to be representative for an average morning rush hour, which means that at certain locations (on a ring road and the city centre) congestion occurs;
- the traffic situation and road network need to allow modelling of temporary local capacity reductions caused by accidents;
- the accidents should have an impact on the traffic situation, without dead-locking the network;
- the accidents should be placed in such a way that a participant is likely to experience the effects;
- the trip of a participant should last around 20 minutes in normal conditions and around 30 minutes in an accident situation (in simulation time).

These requirements are used in the design of the experiment. They are separated in four major elements: (i) road network and Paramics, (ii) scenarios, (iii) participants and (iv) procedure.

7.2.1 Road network and Paramics

For the experiment a fictional road network is used. A fictional network has the advantage that it allows for a a greater experimental control compared to a real road network, as it is possible to control the traffic flows and thus traffic jams, the knowledge of the road network and the traffic conditions. It also means it is possible to control the knowledge and experience of participants of the road network, the traffic situations and possible traffic information. Paramics allows for an easy development of such a fictional road network, including parameters such as the maximum speeds of links, priorities at junctions, route choice algorithms, accidents, etc. This reduces the effort to create a road network, compared to a real road network.

The advantages of using a fictional road network outweigh the disadvantages for this research. The two advantages of a real road network are the use of real photos of junctions and roads and the possibility to compare behaviour in the simulator and in the real word. However, in order to prevent differences between the (perceived) traffic conditions of the real road network and the traffic conditions in the choice simulator, a huge effort would be required to retrieve all the necessary data at the necessary quality. This would be required, as a participant that is familiar with the used road network, will immediately spot any differences in the road network and the photos of intersections. A participant also has a certain familiarity with and perception of network conditions such as location of jams, average speeds, etc. and is likely to use this in the route choices. Such previous knowledge decreases the validity of the outcomes, because it cannot be controlled for.

The hypothetical network means that no photos of existing intersections can be used to display the upcoming junction in the AIDA RCS client. In stead Paramics was used to create a visualization of all entry links for all nodes in the network. These images were created using the 3D viewer of Paramics. However, no buildings or other landscape features of Paramics have been used in these images. There are no landmarks available, although in a real environment, these are often used

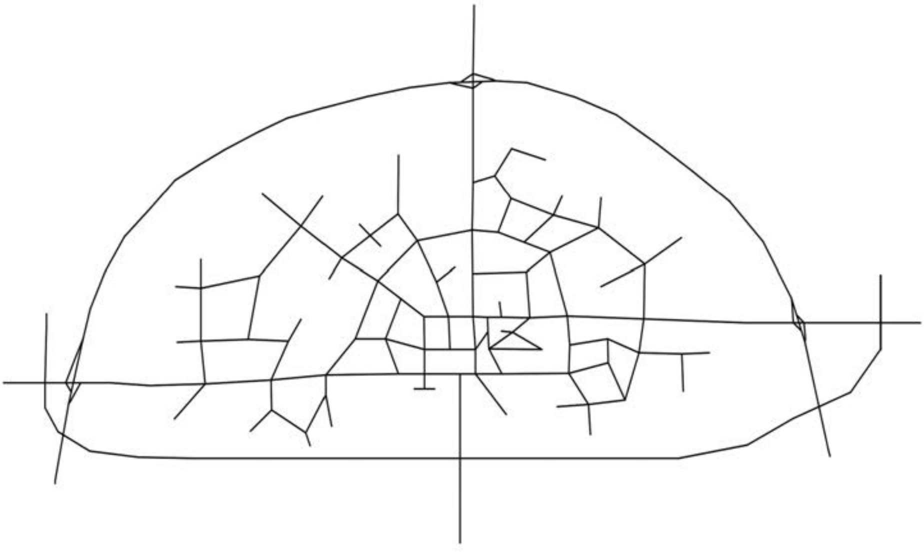


Figure 7.1: *Virtual road network*

for navigation (Burnett, 1998). As no knowledge of the landmarks was available for the participants, it would mean no advantage. The guidance information was not targeted at using landmarks either, which means these are less important. Furthermore, the time need to create a network with a complete city with buildings is not expected to pay off in the results or validity of the experiment.

An overview of the hypothetical road network is given in figure 7.1. The network consists of 58 Origin-Destination zones, 404 links and 175 nodes.

The road network is situated in a city, with a busy city centre in the middle, an inner ring road and an outer ring road, which consists of a highway (northern half) and a provincial road (southern half). As can be seen in figure 7.2, there are four different speed areas. It is easy to recognize the highway and provincial roads, with their respective speeds of 120 km/h and 70 or 80 km/h. There are three areas in the city where the speeds are reduced to 30 km/h. The rest of the road network has an allowed speed of 50 km/h.

Four maps of the road network have been provided to participants before the experiment, to provide them with information on the road network. These are the above mention map of maximum allowed speeds (figure 7.2), a map of the most important street names figure 7.3) and a map with the different neighborhoods (see figure 7.4). The fourth map (not shown) showed the participant his/her origin and destination. All roads in the network have been given a name. For example, the inner ring road is called Ring. The minor roads all have names which match with their appropriate neighborhoods. Participants could position themselves in a certain neighbourhood, but were not able to look up all the individual streets, apart

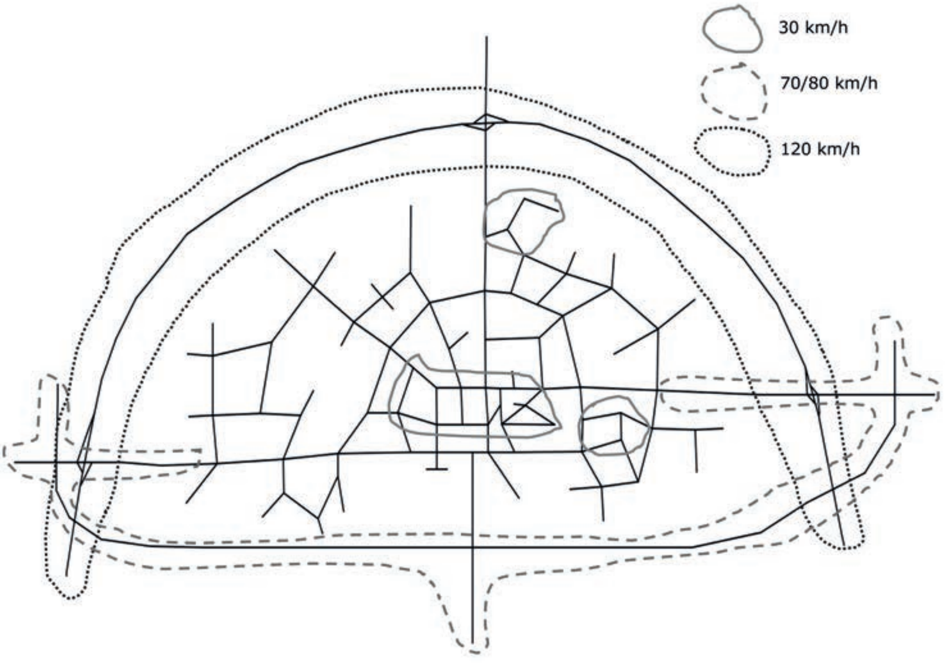


Figure 7.2: *The allowed maximum speeds in the network*

from the major roads. These maps allow a participant to acquire basic network information which is expected to supply the information a participant will need in an unknown road network.

The traffic situation in this network has been created with the tools available in Paramics: the OD matrix, vehicle type distribution and route choice functions. A traffic situation was created in which a limited amount of congestion occurs at logical locations in the road network, such as important junctions on the ring road or exits of the motorway. This congestion is further aggravated as a consequence of accidents. Participants would be likely to encounter the “normal” congestion and/or the consequences of the accidents. To this end, the OD matrix has been tuned, together with the routing algorithm used, to create a traffic situation that matched the targeted situation. Next, the generalized cost algorithm of Paramics for route choice was adapted (see equation (7.1)). Travel time has a much larger importance compared to distance, as it is expected that in a morning rush hour, minimizing travel time is the most important. Distance is still used in order to prevent unreasonably long trips (as this relates to fuel cost).

$$\text{costs} = 2T + 0.5D \tag{7.1}$$

in which:

T = Travel time for link (min)
D = Distance of link (km)

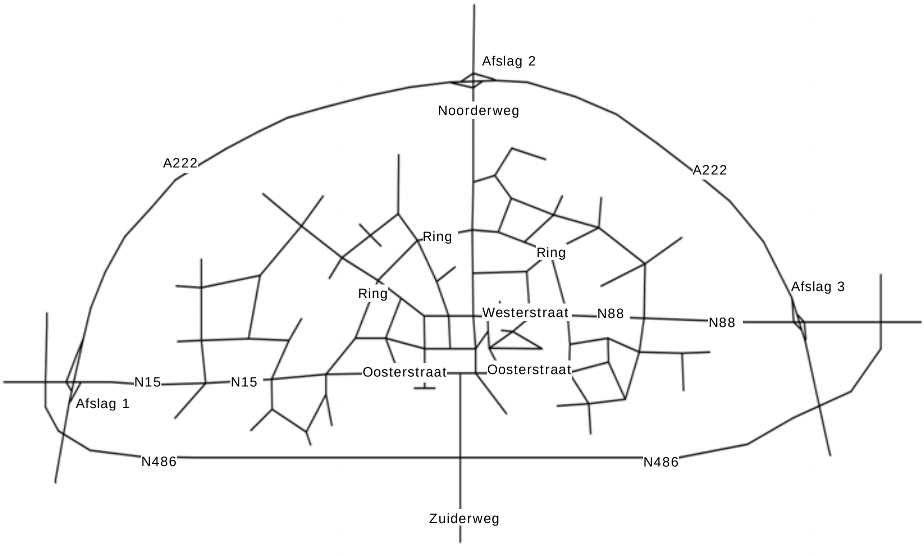


Figure 7.3: *The streetnames of the network (in Dutch)*

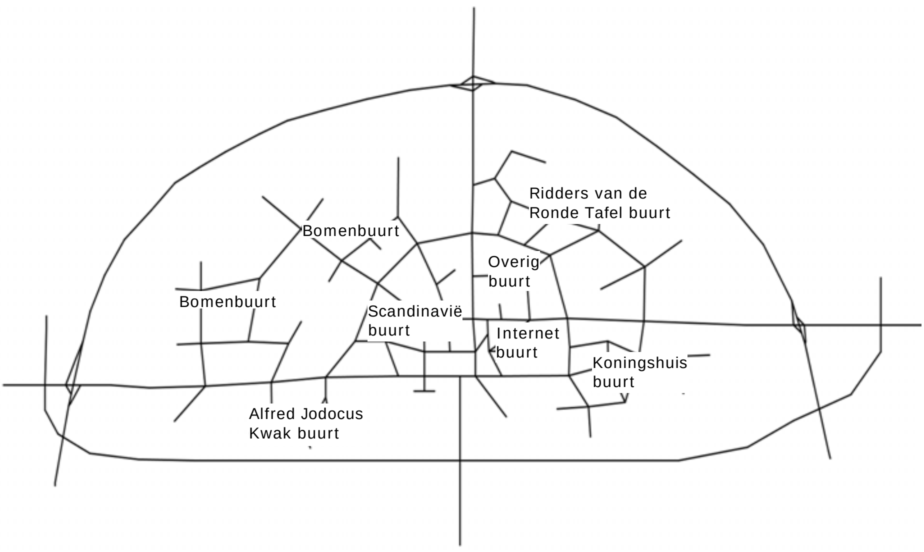


Figure 7.4: *The neighborhoods of the network (in Dutch)*

During the simulation, this function is evaluated every minute, leading to an updated generalized cost for each link, as the actual travel time is part of this cost function. The updated generalized costs are used by familiar vehicles, whereas unfamiliar vehicles will remain using the initially calculated costs. Because the setting for the route choice experiment is a morning commute, a high percentage of familiarity was chosen (85%). Two other factors are also of importance: the feedback smoothing factor and the feedback decay factor. The feedback smoothing factor describes how much of the previous periods is taken into account for the new cost, which is set to 0,8. The feedback decay factor describes how much the costs decays over time if no measurements of travel time are recorded, which was set to 0,995. It only returns to the freeflow travel time. Paramics was also set to use perturbation, which adds a random factor to the total costs. This causes a perceived cost of a trip, in stead of the real costs. The function to calculate the perceived costs is provided in equation (7.2).

$$C' = \left(\frac{100 - P}{100} + N \right) C \tag{7.2}$$

in which:

$$\begin{aligned} C' &= \text{Perceived cost} \\ C &= \text{Original cost} \\ P &= \text{Perturbation factor set to 5} \\ N &= \text{Random number in the range of 0 to 0.1} \end{aligned}$$

After setting the routing algorithm, the demand was tuned between the various origins and destinations to create reasonable traffic situation. It resulted in a static matrix, with a total of just over 36.000 vehicles per hour. This demand matrix is a total over all available vehicles. In total 14 vehicle types have been used, of which around 10% are lorries, 2% coaches and the rest a variation of passenger vehicles.

The virtual starting time of the simulation was set to 7:20 and the maximum duration of the simulation set to two hours. Because in each scenario, the starting time of a participant is around 7:30, the road network has already been warmed up for around 10 minutes, including a build-up of small congestion. The timestep of Paramics describes the number of discrete intervals or time steps per second. It influences the speedup factor (ratio of simulation time to elapsed time) of Paramics, and can be changed during the simulation. During the simulation, the timestep is updated manually by the experiment supervisor, to prevent large deviations of the speedup factor as this factor is influenced by the amount of vehicles in the network.

Regarding the detection of traffic jams, the minimum link length necessary to be taken into account was set to 50 metres, while a traffic jam needed to be more then 1000 metres long before it was to be put forward as information to participants. The update intervals for the regular and semi-regular information were set to 100 milliseconds and 60 seconds respectively. The buffer time was set to 2000 milliseconds and the number of time steps to 5 (refer to equation (6.2) for more information).

Table 7.1: *Network performance*

Variable	No incidents
Average travel time (min)	8:49
Vehicles released	24,189
Average speed (km/h)	66.5
Total km travelled	236,469
Total hours travelled	3,555

The overall network performance of the road network is given in table 7.1, which displays the overall results for a simulation period of 40 minutes, between 7:20 and 8:00. The results for the situation without incidents shows a relative high average speed, mainly because of the large number of vehicles on the highway around the city. Most trip durations are around 9 minutes.

A small amount of congestion occurs at several locations in the network, for example on the Noorderweg joining the Ring, some congestion occurs, as on the Oosterstraat. Some delays also occur on the Ring. The size of the congestion various slightly in this period. This is expected to be a reasonable traffic situation for a normal morning rush hour between 7:30 and 8:00.

7.2.2 Scenarios

In this experiment, each participant will be assigned to a scenario. A scenario is defined here as the sequence of trips a participant will undertake, in which during a trip a specific set of events will take place. As mentioned by Bonsall (2004b), it is important to create a situation for a participant that is both realistic and understandable. The experiment will represent a morning commute, which means that a participant will act as if driving to work, where they are due to arrive at a certain time.

The following restrictions were defined for the development of the scenarios. First, it is important that a participant is able to finish the experiment within 2 hours elapsed time, in order to prevent too much burden on the participant. Second, a participant needs to experience all kinds of traffic information, including no traffic information. There are three kinds of traffic information defined which means that a participant needs to make at least three trips. Third, for each type of traffic information, there must be at least one trip with an accident. Fourth and last, a participant has to have the chance to get acquainted with the AIDA RCS and the way it works.

Based on these restrictions three scenarios have been defined. All three scenarios are a commute trip, starting at around 7:30 (simulation time) and with travel times between 15 and 20 minutes on average, without accidents. The participant is due to arrive at work at 8:00. A speedup factor of 3 was targeted, which means that a trip of 15 minutes simulation time, will take 5 minutes in reality. With around 10 choices per trip, this results in an average of 30 seconds per choice. This means there is enough time available for participants to make a choice, while keeping them concentrated. The duration of the trip in real time also means that respondents

Table 7.2: *Experimental trip setup*

Trip	Scenario	Random Seed	Accident	Traffic information
0	1	2	None	Dynamic route guidance
1	2, 3 or 4	2	None	None
2	2, 3 or 4	4	None	None
3	2, 3 or 4	2	1	None
4	2, 3 or 4	4	2	None
5	2, 3 or 4	2	1	Textual information
6	2, 3 or 4	2	2	Textual information
7	2, 3 or 4	4	1	Dynamic route guidance
8	2, 3 or 4	4	2	Dynamic route guidance

are expected not to experience much experimental fatigue. These three scenarios are used to collect the experimental data.

Before actually starting the data collection a fourth scenario (scenario 1) is used, which was created to let participants get acquainted with the AIDA RCS and the way they should operate. Each participant made 9 trips in total, of which the first was scenario 1 and the remaining 8 trips of one of the three other scenarios. An overview of the total setup is given in table 7.2, which shows that for each scenario two accidents are defined, and two different random seeds for the various models using statistical distributions, which cause a slight difference in the traffic situations.

The order of the trips is the same for all participants. This means no counter balancing or randomization of the order has been performed for the experiment. This was done on purpose, because the effect of any kind of information on the knowledge of the road network, traffic situation or accidents, will influence the remaining trips. For example, if a participant first makes a trip with dynamic route guidance, it becomes clear which route is the fastest, even though an accident has occurred. This knowledge will be used in the next trips. It is assumed that when participants start with no information at first, they will actually consider the information that is presented. This might have an unintended effect of learning during the experiment, which has to be accounted for in the analysis.

All scenarios have different origins and destinations. Scenario 2 has its origin on the highway, scenario 3 on the provincial road, while scenario 4 starts in a suburb. All destinations are in or near the city centre. Each combination of origin and destination has at least two reasonable or logical routes. These routes are more or less equal in travel times. They are also based on the direction of the links and the usage of different link types, such as the highway or ring road. For these two routes, accidents have been defined, which cause a major disturbance of the traffic flows. The three scenarios for data collection can be found in figures 7.5b, 7.6b and 7.7b together with the location of the accidents per scenario.

Each accident consist of certain characteristics. These are presented in table 7.3. Accidents are linked to a certain scenario. Each scenario has a number of routes that are more or less equal, and accidents should be placed on these routes. As the scenarios have different routes, the locations of accidents should also be different. Each accident thus has an associate scenario, linkID and location on that link (not visible in table 7.3). Each accident also has a specific start time, in minutes,

Table 7.3: *Accident setup*

ID	Label	Scenario	Start time (min)	LinkID	Type
1	Accident with 3 cars on A222, between exit 3 and exit 2	2	3	31	Heavy
2	Accident with 2 cars on the Ring, between Noorderweg and Plataanstraat	2	2	209	Simple
3	Accident with 3 cars on the Ring, between Plataanstraat and Buxuslaan	3	4	214	Heavy
4	Accident with 2 cars on A222, between exit 2 and exit 1	3	5	51	Simple
5	Accident with 2 cars on the Ring, between Plataanstraat and Noorderweg	4	4	216	Simple
6	Accident with 3 cars on Oosterstraat, between Laplanderf and Zuiderweg	4	4	160	Heavy

Table 7.4: *Accident types*

ID	Label	Duration (min)	Passing speed (km/h)	Opposing speed (km/h)
Simple	Simple car crash	10	15	20
Heavy	Heavy car crash	25	5	10

which describes the time the accident will happen, once the simulation has started (7:20). Each accident also has a label, which is the text which will be presented to a participant once the accident has happened (thus, once he starts the trip).

Paramics allows each accident to be described in terms of linkID, starting time and location on that link. However, this is not enough to describe the accident. The extra characteristics are given in table 7.4. Wait time describes the duration of the accident. Passing and opposing speed describe the maximum vehicle speeds of vehicles passing the accident (on a different lane) or vehicles on the opposing lanes. Because of the possibility that a respondent selects a route which leads to an accident, each accident is placed on a two lane link. This means that vehicles will still be able to pass the accident location, albeit with a reduced speed. It will still cause congestion, but not completely stop the traffic. A participant will also be able to pass the accident and not experience extreme delays. These locations are chosen to increase the reliability of the experiments.

Placing incidents on the network has effects for all trips, as is shown in table 7.5. In all cases, the average travel time increases, while the average speed decreases. As a results, less kilometers are travelled in the network. All incidents cause extra congestion, in some cases with more impact than in other situations, depending on the location and duration of the accident. For example, incident 1 in scenario 2, causes a traffic jam to extend to the edge of the road network, which heavily influences the average speed of the total network. The largest increase in average travel time occurs with incident 6.

Table 7.5: *Network performance*

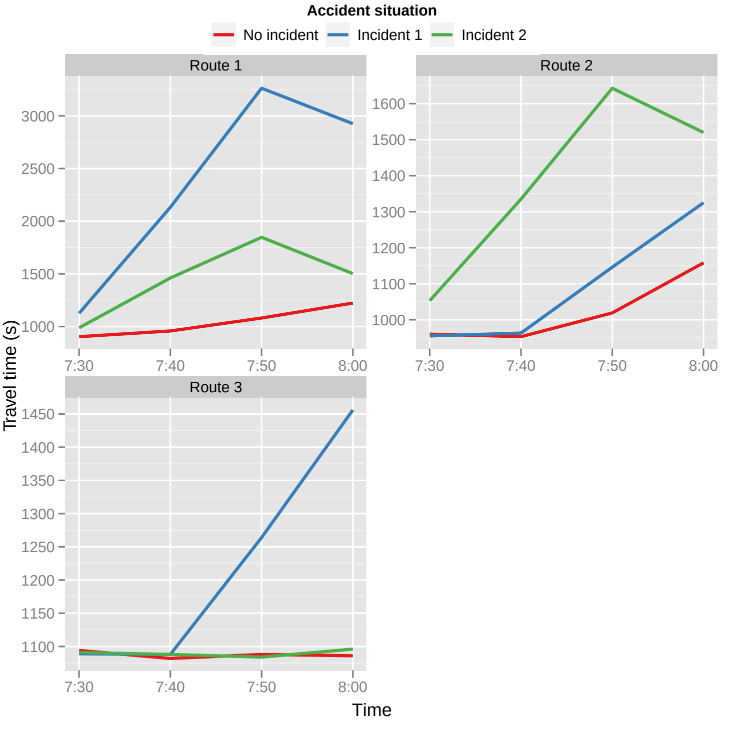
Variable	No incidents	Incident 1	Incident 2	Incident 3	Incident 4	Incident 5	Incident 6
Average travel time (min)	8:49	9:21	9:09	9:03	9:11	9:07	9:25
Vehicles released	24,189	23,932	24,141	24,224	24,237	24,237	24,175
Average speed (km/h)	66.5	55.9	63.1	64.7	61.7	63.9	61.4
Total km travelled	236,469	208,498	232,384	236,348	228,788	234,914	232,874
Total hours travelled	3,555	3,371	3,684	3,652	3,709	3,679	3,793

For each of the three scenarios of interest and all the accidents within these scenarios an investigation has been made on the travel time of the most reasonable routes at different time instances, between the origin and destination of that scenario.

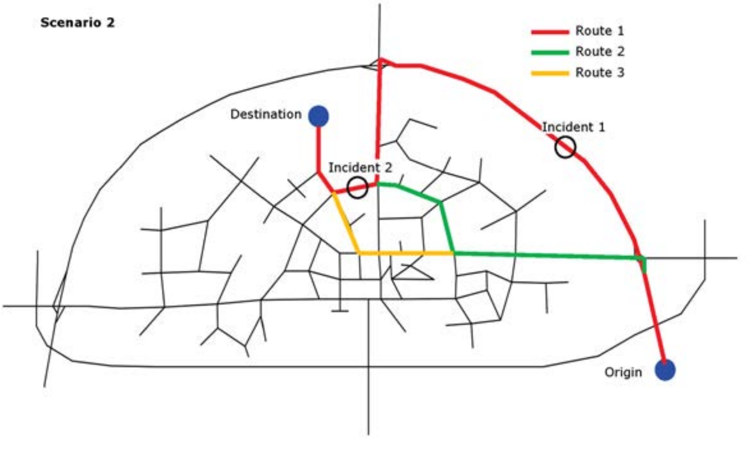
For the travel times in scenario 2 (see figure 7.5a) it is clear that the location of the incident has a large influence for a route that runs past this incident location. For example, incident 1 is located on route 1, which results in a travel time at 7:50 of more than an hour. This is more than three times the travel time without an incident on that route. Incident 1 also shows an effect on the travel time for all routes, because the location of the accident causes congestion to push backwards onto these routes, before the exit on the highway. The location of incident 2 has the most effect on route 2, and slightly on route 1, because both routes use the link which has this accident. Of these three routes, route 3 only shows an effect with incident 1, and does not experience much effects of rerouting traffic because of the other incident.

For route 4 it is clearly visible that incident 4 has a large effect on its travel time (see figure 7.6a), although the effects are not as large as for route 1 and incident 1 in scenario 2. The difference is caused by the location of the incident, not close to the edge of the network. Routes 5 and 6 show no big differences for incident 4, although route 5 shows an upwards trend between 7:50 and 8:00, which is most likely caused by rerouting traffic avoiding the highway. Incident 3 has no effect on the travel times for route 4, but a large impact on route 5, because the incident is directly on this route. This incident shows a larger effect for route 6, as the travel time increases more than in the other situations. This is not necessarily due to congestion blocking the entry of this route, but by rerouting traffic. Without accidents, the travel time for route 6 also increases, because of the amount of traffic moving through the city center.

Scenario 4 has two reasonable routes which are expected to be used by participants (see figure 7.7b). The resulting travel time for these routes are given in figure 7.7a. Without accidents, both routes show a small increase of travel time at the start of the simulation, but are relatively stable for the rest. Route 8 is slightly faster than route 7, but just with a margin of around 30 seconds. This is not very likely to be perceived by participants. For both situations with incidents, the effects on both routes are clearly visible with marginal effects on the other route. Route 7 runs over the location of incident 5, whereas route 8 passes the location of incident 6.

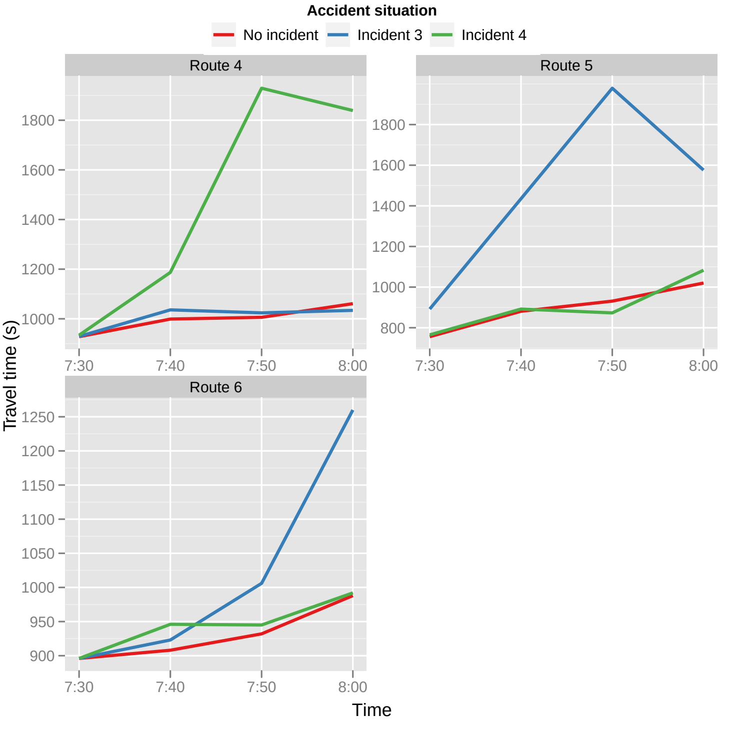


(a) *Travel time*

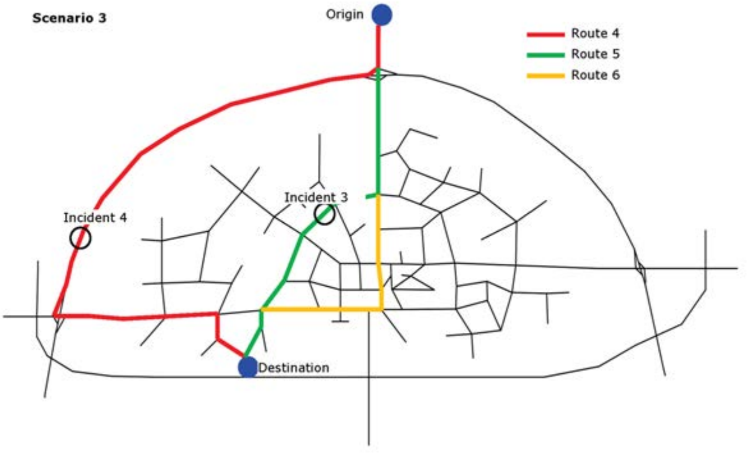


(b) *Possible routes*

Figure 7.5: *The travel time of the three routes for scenario 2*

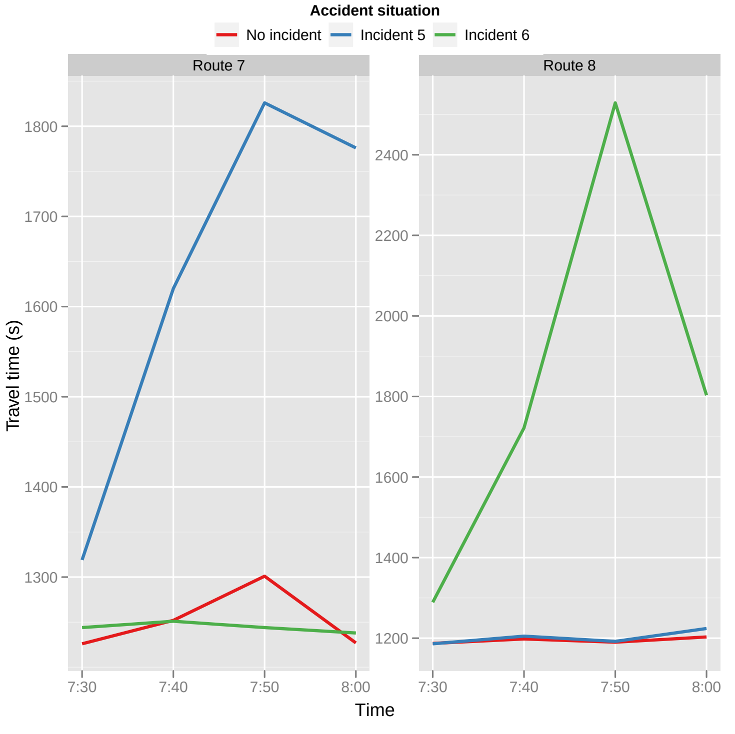


(a) *Travel time*

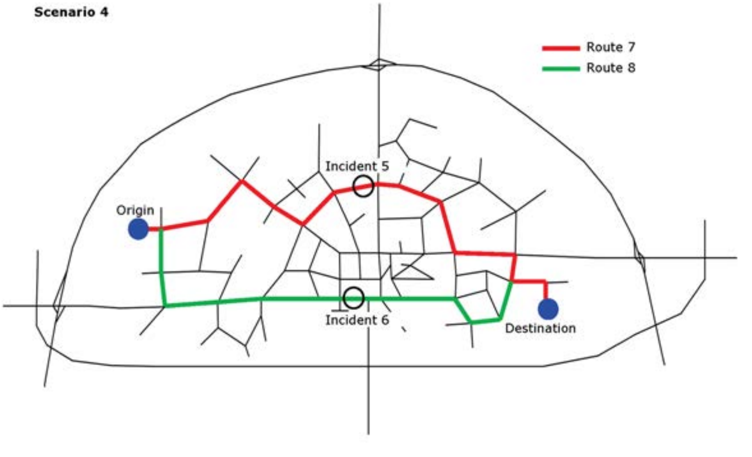


(b) *Possible routes*

Figure 7.6: *The travel time of the three routes for scenario 3*



(a) *Travel time*



(b) *Possible routes*

Figure 7.7: *The travel time of the three routes for scenario 4*

7.2.3 Participants

The participants for the experiment were selected amongst colleagues, students and friends. Because of time constraints for the actual experiment, no explicit sampling strategy has been used. Participants did not receive any fee for taking part in the experiment. For each scenario, ten participants are expected to provide for enough data to allow various analysis. With an estimated average number of twelve choices per trip, this means a total of 2880 choices at junctions over all trips, scenarios and participants. For each combination of traffic information and accident situation, 10 complete trips are available, and an estimated 120 choices. However, the total amount of 30 participants means it is not possible to achieve an appropriate representation of Dutch car drivers.

7.2.4 Procedure

The AIDA RCS was conducted with at least two computers. One which was running Paramics, the database and the RCS Server. The other computers were connected to this computer, and ran the RCS client. It was made sure participants did not see the screen with Paramics, as this would have influenced their choices and trips. Each participant was given a introduction to the experiment (see appendix D), and was given some time to ask questions. The participant also received four maps of the road network (OD map, speed map, neighborhood map and street map) and a table in which all the settings a participant had to enter in the RCS client (type of traffic information and scenario). During the first trip, participants could ask additional questions. For the 8 trips of interest (see table 7.2), participants were left to focus on the experiment. Of each trip, apart from the database logging, also a log with the exact ID's of simulation and trips was made for each participant. This log serves as a check for the database logging and the experiment and problems during trips, as Paramics occasionally halted during the trips. After finishing each trip, the participant answered the post-trip questions. Once all participants finished their trips and questions, Paramics and the RCS server were configured for the next trip. The total duration of the experiment for one participant is between 2 to 2.5 hours.

7.3 Summary

The AIDA RCS is used to facilitate an experiment to acquire route choice data in situations with accidents where participants received traffic information. The trips are made in a hypothetical road network, to certify that all participants started with the same level of knowledge of the road network and traffic conditions. Each participant (30 in total) was assigned to one of three scenarios, where each scenario consisted of 9 trips. The first trip was used to let participants get accustomed to the AIDA RCS, where the following 8 trips all had the same order of types of traffic information. The scenarios differed with respect to the origin, destination and locations of the accidents. Before the scenario, each participant had to provide personal data, and after each trip the participant answered three questions regarding the level of realism and route choice. During each trip, the choices are logged in the database, together with information on the traffic situation in general, and the specific traffic information and/or route guidance a participant has received. This data enables us to investigate effects of both incidents and traffic information on the route choice in general, as well as the estimation of various route choice models.

Success is not final, failure is not fatal: it is the courage to continue that counts.

— Winston Churchill

Chapter 8

Route choice experiment - General results

8.1 Introduction

This chapter presents the results of the route choice experiment using the AIDA Route Choice Simulator (RCS). The results of the experiment are used to provide an answer to research question 3: *What are the effects of traffic information on route choice during non-recurrent traffic situations?* Within this research question, a number of items are of specific interest. These are:

- the realism of the AIDA RCS and the experiment;
- development of a preference for a route;
- differences in route choices in case of incidents and guidance;
- the effects of incidents on the characteristics of the chosen routes in terms of length, travel time, delay and number of routes;
- the effects of traffic information and/or route guidance on the characteristics of the chosen routes in terms of length, travel time, delay and number of routes.

The chapter starts with an overview of the participants who took part in the experiment (section 8.2), followed by the results of the self-reported realism by the participants (section 8.3). Section 8.4 presents the routes chosen by the participants for various situations, using graphical overviews and amount of routes and choices for these situations. The characteristics of the chosen routes and the effects of incidents and AITS are presented in section 8.5 followed by a discussion of the AIDA RCS, the experiment and the results in section 8.6. The chapter finishes with a summary in section 8.7.

8.2 Participants

An overview of the characteristics of the participants is presented in table 8.1. In total 33 participants took part in the experiment, of which the data of three participants could not be used, due to technical malfunctioning. This resulted in a total of 30 participants of which the data is used in the analysis.

Most participants were male, between 22 and 30, with a Master title, possessed their drivers license between 5 and 15 years, and drove between 0 and 5000 km per year. In other words, most of the participants were young professionals who recently graduated. This means, when compared to the internet survey, the difference with the total Dutch driver population is larger.

Table 8.1: *Participant overview*

Variable		**Amount**	**%**
Gender	Male	21	70%
	Female	9	30%
Age	18-21 years	1	3,3%
	22-30 years	15	50.0%
	31-40 years	7	23,3%
	41-50 years	1	3,3%
	51-60 years	5	16,7%
	61-70 years	1	3,3%
Education	High school (MAVO)	1	3,3
	High school (VWO)	1	3,3%
	HBO	6	20%
	University, Bachelor	3	10.0%
	University, Master	19	63,3%
Drivers license	0-2 years	3	10.0%
	2-5 years	2	6,7%
	5-10 years	8	26,7%
	10-15 years	8	26,7%
	15 years or more	9	30.0%
Kilometers per year	0-5.000 km	13	43,3%
	5.000-10.000 km	6	20.0%
	10.000-15.000 km	3	10.0%
	15.000-20.000 km	2	6,7%
	20.000-30.000 km	3	10.0%
	30.000 km or more	3	10.0%

8.3 Self-reported realism

After each completing a trip, each participant was asked three questions about their experiences of that specific trip. Two of these questions are related to the realism of the AIDA RCS and the experiment. These are indicators of the validity of the experiment, as perceived by the participants. These questions are:

1. What was your general experience of this trip?
2. Would you make the same route choice in real life?

The results for the first question on the general experience is given in table 8.2. A total of five responses are missing (because of an undetected bug at the start of the experiment) although the participant did finish the trip, the AIDA RCS did not present the post trip questions to the participants.

The responses indicate that most participants report a realistic or very realistic trip experience (78,1%), whereas only 14,5% indicate the trip to be (very) unrealistic. The type of ATIS that was available in each trip had no significant influence on the realism. Scenarios however did have a significant impact on the reported realism, as in scenario 4, more participants said the general experience was very realistic. Scenario 3 has the lowest reported realism, but still has most participants experiencing each trip as realistic. Specific incidents have a small impact on the realism, and as these are related to the scenarios the responses are significantly different between the incidents.

The next questions asked participants if they would make the same route choice in a real life situation. The results are given in table 8.3. Again 5 answers are missing.

Table 8.2: *Answers for the question "What is your general experience of this trip?" per incident situation*

		Answer						
Scenario	**Incident situation**	Very realistic	Realistic	Not realistic	Very unrealistic	No opinion	I had troubles during the trip	Missing
2	No incident	20%	65%	5%	0%	10%	0%	0%
	Incident 1	23%	57%	10%	0%	10%	0%	0%
	Incident 2	10%	63%	17%	0%	7%	0%	3%
3	No incident	0%	60%	30%	0%	0%	10%	0%
	Incident 3	13%	67%	17%	0%	0%	0%	0%
	Incident 4	10%	70%	13%	0%	0%	0%	7%
4	No incident	40%	40%	20%	0%	0%	0%	0%
	Incident 5	37%	50%	10%	3%	0%	0%	0%
	Incident 6	40%	37%	10%	0%	0%	7%	7%
Total		22%	57%	14%	0%	3%	2%	2%

Table 8.3: *Answers for the question "Would you make the same route choice in real life?" per incident situation*

		Answer			
Scenario	**Incident situation**	Yes, certainly	Yes, possibly	No	Missing
2	No incident	40%	55%	5%	0%
	Incident 1	67%	27%	7%	0%
	Incident 2	47%	33%	17%	3%
3	No incident	40%	20%	40%	0%
	Incident 3	37%	50%	10%	3%
	Incident 4	30%	50%	13%	7%
4	No incident	30%	65%	5%	0%
	Incident 5	37%	53%	10%	0%
	Incident 6	47%	50%	7%	7%
Total		40%	45%	13%	2%

In 40% of the trips made, the participants responded they would certainly make the same route choice in real life. 45% said they would possibly make an equal route choice, which results in 85% of the trips which are likely to have equal route choices in real life. Scenario 2 has the highest amount of responses for a certain equal route choice, whereas scenario 3 has the largest amount of responses saying they would not make such a route choice. This difference is significant ($p<0.05$). The type of traffic information caused no significant difference between the responses, although in case of dynamic route guidance, most participants were certain about a similar route choice. Incidents neither caused a significant difference in the responses, although some variation exists, again for scenario 3.

Combining the results of both questions, in total 35 trips were made which were said to be unrealistic and 29 trips where the same route choice will not be made in real life (out of a total of 240 trips). Of these trips 13 trips combine a low level of realism and a different route choice, which can be shown on a map to be "weird" route choices, which are not likely to be made in real life. Of these trips, 8 were made by two participants. This indicates two of the thirty participants had more than average difficulty in taking part in the experiment in the AIDA RCS.

In general, the participants reported to have a realistic experience with the experiment in the AIDA RCS. They said the trips were (very) realistic and they would certainly make an equal route choice in real life. The face validity of the experiment and the AIDA RCS are similar to previous findings on face validity of route choice simulators in general. However, as the questions did not use a strict definition of "general experience" and "route choice in real life", the results depend on the perception of the participants on these terms. It is not expected to cause large differences, and as such will not have a significant impact on the face validity.

8.4 Chosen routes

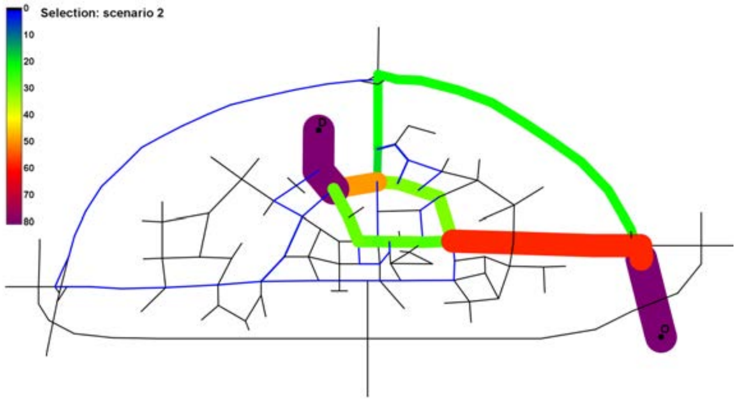
This section provides an overview of the routes that participants chose in the AIDA RCS. The routes are depicted using maps, and are analysed using the counts of different routes. The overviews are presented for all scenarios, and for the variations with type of guidance and incident situations. The section starts with an analysis of the "preferred" routes in section 8.4.1 in relation to the "reasonable" routes. The differences which are caused by incidents and type of guidance are presented next, in section 8.4.2.

8.4.1 Preferred routes

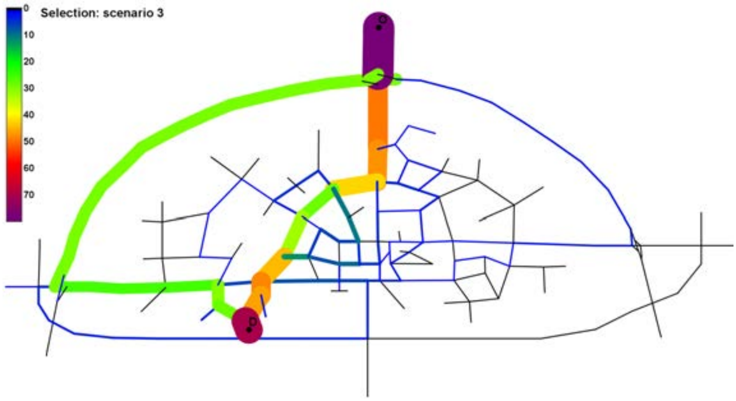
The routes chosen by each participant in each scenario are pictured in figure 8.1. These overviews consist of all trips made, including all variations with and without incidents and the various types of information. In scenario 2, it is obvious that three main routes were chosen by participants, one of which consisted mainly of highway and two others used roads in the city, of which one actually went straight through the city centre. In a few cases, participants made other choices, perhaps because they mixed right and left turns at intersections. Scenario 3 shows a different pattern, where two obvious routes are chosen most often. One of these routes is the highway, while the other uses the ring road of the city. The variation in routes around the ring road are possibly due to either mixed right and left turns, or choices for routes that avoid the effects of the incidents themselves. Scenario 4 shows much more variation in the most chosen routes. Two of them can be considered to be the main alternatives: either the southern route using the Oosterstraat, or the northern route using the ring road. However, various alternatives have been used, possibly to avoid congestion because of incidents or the incidents themselves.

In all three scenarios, these maps indicate that participants were able to find a "preferred" route. These preferred routes are a close match with the predetermined "reasonable" routes in figures 7.5b, 7.6b and 7.7b. Scenario 2 shows the highest overlap with these routes, whereas for scenario 3 one of the "reasonable" routes (heading south using the Noorderweg) is not perceived by the participants as such. Of the other "reasonable" routes, the one using the ring road is likely to be preferred most. The map of scenario 4 shows that the "reasonable" route using the Oosterstraat is preferred over the other route, but participants did have a larger difference in the preferred routes here. These results are confirmed by the popularity of the different routes per scenario (see table 8.4). This table shows the amount of times a route is used; for example 31 times for one route or 1 time for five routes in scenario 2.

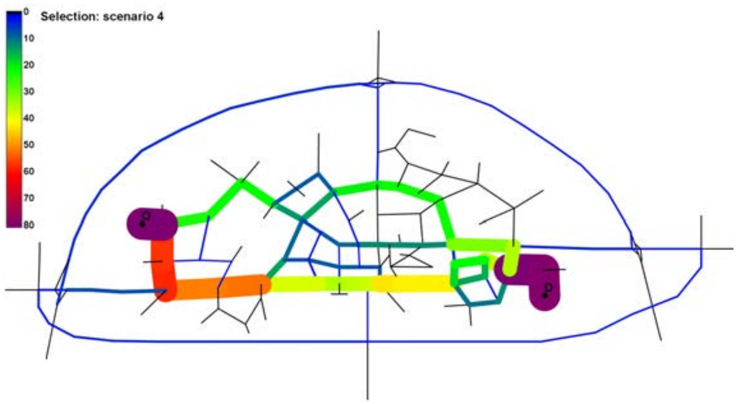
The results for scenario 2 clearly show that three routes are used most, with 5 different routes being driven just once. This matches the results of figure 8.1a and indicates that the participants have a clear preference for three routes. Scenario 3 shows a larger variation in number of routes selected. In this case, only two routes show a large number of uses, whereas 19 other routes are only used once. Two routes in this scenario are preferred, but participants show a lot of variation in route selection (as can be seen in figure 8.1b). This variation is most likely due to



(a) *Scenario 2*



(b) *Scenario 3*



(c) *Scenario 4*

Figure 8.1: *Overview of used routes per scenario (color indicates number of trips using a link)*

Table 8.4: *Number of routes and their popularity per scenario*

Scenario 2		**Scenario 3**		**Scenario 4**	
Popularity of a route	*Number of routes*	*Popularity of a route*	*Number of routes*	*Popularity of a route*	*Number of routes*
31	1	23	1	13	1
26	1	22	1	11	1
18	1	5	1	7	1
1	5	3	2	6	2
		2	2	5	1
		1	19	4	1
				2	6
				1	16

the orientation of the map in the AIDA RCS client and the position of the origin and destination (which was mentioned by participants as a reason for wrong choices during the experiment). Scenario 4 is different from the other two scenarios, as the number of times the most often routes are used, is lower (13 and 11 times for scenario 4, compared to 23 and 22 for scenario 3 and 31 and 26 times for scenario 2). A likely reason is the availability of a larger number of viable and equivalent alternatives which causes participants to differ in their route choice. One of the areas where participants clearly had differed in their "preferred" routes is the last part, when they had to choose their route through the Koningshuis buurt, or use a part of the N88. Figure 8.1c clearly show the variation of routes.

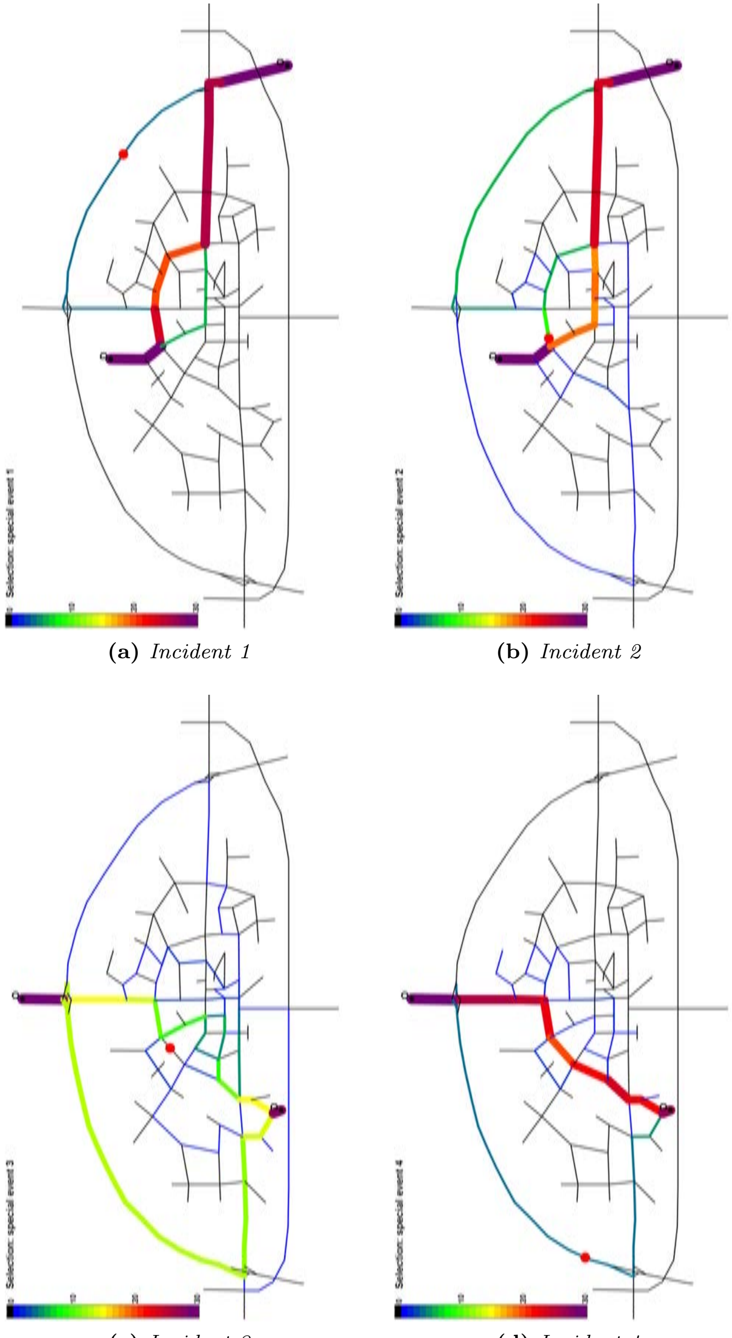
The diagrams suggest that most participants have determined a preferred route. Even with only 8 trips, participants are able to select a route which suits their personal preferences. The preference coincides with the predetermined "reasonable" routes. This is confirmed by the amount of routes used by a particular participant. Some only use 2 or 3 routes for these 8 trips, with an average of 5 routes used per participant. The variation in the routes used per participant are largest for scenario 3, which is confirmed by the graphic overviews.

8.4.2 Effects of incidents and guidance on routes

Whereas in section 8.4.1 the aggregated routes are presented, this section concerns the routes under different circumstances. Apart from the scenarios, incidents and type of guidance varied. This section presents the findings on the effects that incidents and type of guidance have on the chosen routes.

The routes which are chosen for the different incident situations are given in figure 8.2. The results are aggregated over the type of guidance that was available for the participants, so in total 30 trips are available for each incident situation. The specific incident locations are shown with a red dot on the map. Once a participant started the trip, the available traffic information on incidents or congestion were displayed on the windscreen (in case traffic information or dynamic route guidance were available).

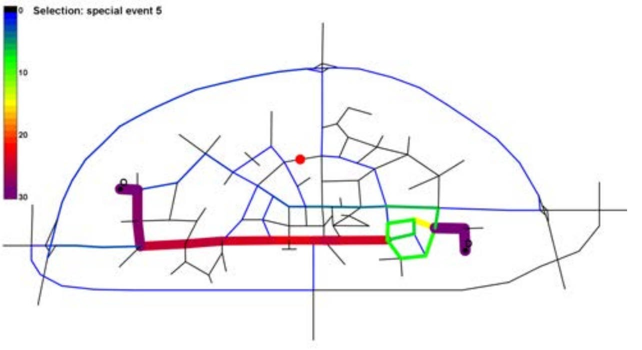
For the incidents in scenario 2 (figures 8.2a and 8.2b), the impact of the incident is obvious. The incident on the highway means most participants selected the ring



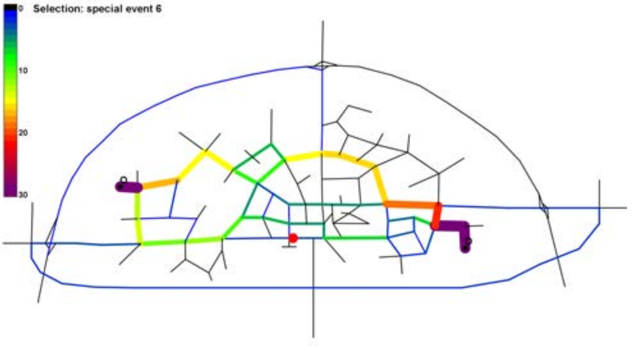
(a) *Incident 1* **(b)** *Incident 2*

(c) *Incident 3* **(d)** *Incident 4*

Figure 8.2: *Overview of most used routes per incident; a red dot denotes the incident location (color indicates number of trips using a link)*



(e) *Incident 5*



(f) *Incident 6*

Figure 8.2: *Overview of most used routes per incident; a red dot denotes the incident location (color indicates number of trips using a link) - continued*

road (only 5 trips use the highway), whereas the incident on the ring road shows more participants select the route through the city centre. Less than 10 trips pass this incident location.

For the incident on the ring road of scenario 3 (figure 8.2c) the first choice is made at the on ramp for the highway. The amount of trips using either is almost the same (a difference of 2 trips). However, the route through the city shows quite some detours, once participants realize they are running into an accident and its consequences. A number of the chosen routes are also completely out of the directions that can be expected. This may be due to the difficulty participants had with orientation during their trip from north to south. In case of the incident on the highway (figure 8.2d), it is obvious that the preferred route is the one using the ring road, with only five trips using the highway. Again a few routes indicate that participants had trouble in selecting the "right" exit at an intersection.

The most used routes for the incidents in scenario 4 are shown in figures 8.2e and 8.2f. The influence of the location of incidents in this scenario is striking. The most used route for the incident on the ring road is very much alike for most trips (with over 25 trips using this route), avoiding the incident from the start, and using

Table 8.5: *Average number of routes used per participant per incident situation, weighed by the number of trips made for a combination of scenario and accident situation*

Scenario	Accident situation	Average number of routes	Weighed average
2	No incident	1.5	1.13
	Incident 1	1.9	1.20
	Incident 2	2	1.33
3	No incident	1.9	1.81
	Incident 3	2.5	2.08
	Incident 4	2.1	1.47
4	No incident	1.8	1.62
	Incident 5	2.4	1.92
	Incident 6	2.8	2.61

the main arterial (the Oosterstraat) through the city centre in the south. None of the participants actually traversed the link with the incident. For the incident on the Oosterstraat, the variation between the used routes is much larger. More trips start with another first choice, and most of these prefer the ring road for the rest of their journey. Of those that pick the southern side of the city, most reroute through the city centre, avoiding the incident itself.

These maps show that the incidents have a large impact on the chosen routes. Based on these maps the routes found for incident 1, 4 and 5 are most likely to be the "preferred" routes, as the number of routes with a high usage are found for these situations. The effects found for the other incidents are easily explained using these "preferred" routes, as all three situations the incident is located on this route. This causes participants to differ in the routes chosen.

This result is confirmed by the average number of routes a participant used (see table 8.5). Herein, the average number of routes is weighed to compare the averages for the trips with and without incidents. Incidents 1, 4 and 5 have the lowest weighed average number of routes, compared to the other trips with incidents. This is most clear for incidents 3 and 6, which is the same as found in figures 8.2c and 8.2f. This indicates that the location (and effects) of these incidents cause participants to re-evaluate their route choices and try to improve their routes. Apart from scenario 3, it also shows that incidents have an effect on the number of routes used. In case of incident 4, this is likely due to the fact that participants learned how to navigate in the AIDA RCS with the north-south orientation, together with the chosen route to be really the best route.

For the effect of the different types of ATIS used, the popularity of different routes is given in table 8.6. Compared to the overall route popularity given in table 8.4, the number of times a route is used is lower for all types of ATIS. Part of this is due to the fact that the results for the different types of ATIS are aggregated over all scenarios with different routes. The differences between textual information and dynamic route guidance indicate there is much more agreement in the routes used when dynamic route guidance is available, especially for the routes that are used only once. Dynamic route guidance influences the routes chosen, and is likely to make participants select the same (advised) routes. Without dynamic route guidance,

Table 8.6: *Number of routes chosen per type of guidance*

Type of guidance	Popularity of a route	Number of routes
No information	17	1
	14	1
	12	1
	9	1
	8	1
	7	2
	4	1
	3	1
	2	3
	1	32
Textual information	10	1
	8	1
	7	1
	5	1
	3	1
	2	5
	1	17
Dynamic route guidance	9	2
	6	1
	5	2
	3	3
	2	4
	1	7

there is much more difference between the used routes, and thus the amount of used routes increases. No information again has a larger variation in chosen routes, also due to the larger amount of trips made without information. Both textual information and dynamic route guidance however cause a concentration effect.

The same results are presented for the different types of traffic guidance, combined with the three scenarios (see table 8.7). In this case each participant had four trips with no information and two trips with both textual information and dynamic route guidance. These results show that participants used two different routes in most cases for dynamic route guidance and, apart from scenario 2, also for textual information. This indicates that participants change their route based on the information they receive about congestion and/or accidents, and the guidance they receive. Scenario 2 is a little different from the other scenarios, as the average number of routes for no information and textual information is different from the other scenarios. In this scenario, the lack of information clearly is of no problem for participants, as they hardly have to vary their route, even with an incident. Textual information also has no effect on the weighed average number of routes, while route guidance makes them use more routes on average, comparable to the other scenarios. For textual information and dynamic route guidance, each scenario had two incidents, with for each incident a different fastest route. If the participants would follow the route guidance advice in these scenarios, the average number of routes would be 2. This only happens in scenario 4, which means that in the other scenarios some participants chose another route than the advice provided by dynamic route guidance. Still, the concentration effect found for the total number of routes used can be confirmed with the average number of routes per participant.

Table 8.7: *Average number of routes used per participant per type of traffic guidance, weighed by the number of trips made for a combination of scenario and type of guidance*

Scenario	Type of guidance	Average number of routes	Weighed average
2	No information	2.1	1.10
	Textual information	1.4	0.98
	Dynamic route guidance	1.9	1.81
3	No information	3.2	2.56
	Textual information	2	2
	Dynamic route guidance	1.9	1.81
4	No information	3.4	2.89
	Textual information	2	2
	Dynamic route guidance	2	2

In general, both incidents and ATIS have an impact on the routes that are chosen by the participants. The location of incidents is very important. If a location is located on the "preferred" route, participants show a larger variation in their chosen routes as they try to find a route without effects caused by the incident. An incident which is not located on such a "preferred" route, hardly has an effect on the chosen routes. ATIS allows for the participants to select a route which is fastest. As such, they know early on if they need to chose for a route differing from their "preferred" route. Especially dynamic route guidance also shows concentration effects, as participants are likely to comply with the advice provided.

8.5 Route characteristics

All the routes chosen by the participants can be described using their characteristics in terms of length, (freeflow) travel time and delay. These characteristics are discussed here, with respect to the different scenarios, incidents and types of guidance. This section starts with the actual characteristics in section 8.5.1, followed by a more in depth discussion of travel time and delay in section 8.5.2.

8.5.1 Overview of route characteristics

Table 8.8 shows the information of the total travel time, freeflow travel time, delay and trip length for each combination of scenario, type of information and accident situation, aggregated over all trips.

In the three scenarios, the average travel time (without incidents and information) is between 18 and 24 minutes, with an average delay between two and five minutes. The overall reduction of travel time for the trips with incidents is 13,45% and 19,25% respectively for textual information and dynamic route guidance. Providing information, regardless of the incident and scenario, thus has a rather large influence on the travel time the participants experienced. The reduction of the travel time is

Table 8.8: *Average and standard deviation of route characteristics, per scenario, type of guidance and accident situation.*

Scenario	Type of guidance	Accident situation	Travel time (s)		Freeflow travel time (s)		Delay (s)		Length (km)	
			Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
2	No info	No incident	1095.9	166.9	824.4	99.3	271.5	182.2	18.5	2.5
		Incident 1	1414.8	589.3	819.6	58.9	595.2	637.2	17.4	2.1
		Incident 2	1328.4	314.9	925.5	216.7	402.9	274.0	17.9	4.1
	Textual info	Incident 1	969.4	41.7	883.1	41.5	86.3	19.2	15.7	0.0
		Incident 2	1179.5	211.9	935.7	134.7	243.8	213.8	16.8	2.3
	Dynamic Route Guidance	Incident 1	962.1	32.3	857.7	25.4	104.4	32.3	15.7	0.0
		Incident 2	1136.3	282.8	1028.5	191.4	107.9	93.5	19.0	7.3
3	No info	No incident	1199.1	657.7	978.0	516.6	221.2	162.1	22.1	12.4
		Incident 3	1289.2	364.3	1011.7	208.7	277.5	235.7	20.7	6.7
		Incident 4	967.3	229.9	762.0	181.7	205.3	96.9	15.2	5.6
	Textual info	Incident 3	1178.0	250.0	930.5	94.4	247.5	241.3	18.9	3.7
		Incident 4	906.7	70.7	714.5	76.0	192.2	50.1	12.3	0.9
	Dynamic Route Guidance	Incident 3	1073.6	195.6	897.8	147.0	175.8	128.2	17.7	4.2
		Incident 4	858.1	112.9	698.5	67.3	159.6	61.7	17.7	4.2
4	No info	No incident	1404.3	214.5	1226.3	159.9	178.0	104.3	18.0	4.5
		Incident 5	1476.5	183.4	1246.2	155.2	230.3	99.4	20.2	6.9
		Incident 6	1640.3	318.0	1443.3	288.6	197.0	203.1	22.5	8.9
	Textual info	Incident 5	1360.2	140.1	1203.0	137.3	157.2	35.4	17.4	3.2
		Incident 6	1436.7	236.0	1220.4	97.7	216.3	198.5	18.5	4.1
	Dynamic Route Guidance	Incident 5	1247.1	45.3	1104.2	27.5	142.9	48.7	16.4	0.3
		Incident 6	1282.0	121.9	1160.0	57.7	122.0	72.6	16.6	0.8

significant (t-test with $p<0.05$) for both textual information and dynamic route guidance compared to no information. The addition of route guidance to the textual information does not significantly reduce the travel time (t-test with $p>0.05$).

The average delay over all scenarios, incidents and types of traffic information is around 3.5 minutes, which is 17,8% of the total travel time. Without traffic information, the average delay over all scenarios and incidents is over 5 minutes. Traffic information and dynamic route guidance reduce the average amount of delay with 2 minutes and 3 minutes respectively. This is a 40.2% reduction with just textual information and 57.5% reduction with textual information and dynamic route guidance. Both reductions are statistically significant ($p<0.05$), and the addition of route guidance to textual information also is significant. The reduction of travel time is mainly caused by the reduction in delays. 70.3% of the reduction in travel time is caused by a reduction of the delays when textual information is provided and 70.1% of the reduction of travel time between textual information and dynamic route guidance is caused by a reduced delay. The other cause of reduction is a change in free-flow travel times. Traffic information causes a reduction of around 1 minute and dynamic route guidance further reduces the free-flow travel time by 20 seconds. ATIS in general thus allow participants to reduce the delay experienced by incidents and congestion, but also improve the route with respect to the free-flow travel time.

For the trips where textual information has been provided, it is clear that the average trip length is decreased compared to the situation without information (but with incidents). Also the standard deviation decreases. Participants are more likely to be in agreement, as already shown in section 8.4. The same decrease in average trip length and standard deviation of this length can be found when dynamic route guidance is provided. Especially the standard deviation is decreased to close to zero, which means that there is hardly any difference left between the used routes of all the participants. This coincides with the results for the freeflow travel time, where a reduction of the standard deviation means more overlap in the different routes chosen by participants.

The results are further detailed for the various scenarios and incidents. In scenario 2, 25% of the travel time is caused by delay aggregated over all trips. For scenarios 3 and 4, this is respectively, 18% and 11%. These percentages are quite large. Scenario 3 also has a very large standard deviation for the travel time, which is almost half the average. This indicates a lot of variation in the routes and thus in the travel times of the participants. This matches the large standard deviation of the trip length in this scenario. The most plausible reason for this large variation is the problems participants had with having a route from north to south through the road network, which made it hard for some participants to match right and left turns.

For the situations without traffic information with incidents, scenario 2 and 4 both show in increase of average travel time and in most cases an increase of standard deviation. It is clear that without information, participants experience the effects of incidents, as they were unable to select a route without incidents beforehand. As a result, the delay is also increased, together with its standard deviation. However, the freeflow travel time in scenario 2 has not increased as much as the delay, which shows that participants used the same route as before. This results in a travel time with a larger delay (between 30% and 42%). For scenario 4, the percentage of the travel time consisting of delay is hardly increased. This indicates that participants were able to select other routes around the incident, which increased their total travel time, but not their delay.

Scenario 3 is different in this case, because for one of the incidents, the average travel time and the standard deviation is much lower than without an incidents. This effect was also found for the trip length. The most likely cause is a learning effect, because participants are better able to make choices for their route and distinguishing right from left when moving from north to south. This especially applies for incident 4, where the average travel time is significantly lower compared to the situation without incidents. The location of the incident is in this case not on the most preferred route of these participants, while incident 3 is, because there is an increase in average travel time and in delay for this situation. The fact that the standard deviation of the freeflow travel time is also decreased, means that there is much more agreement in the routes that participants used. This is also confirmed by the routes that are depicted in figure 8.2d.

Regardless of the scenario, the start time of the participants varied based on a combination of Paramics releasing the controlled vehicle and the participant pressing the spacebar. The actual start time of the participants were within a minute of the proposed start time of 7:30. As this variation could have influenced the information available as well as the advised route, the correlation with travel time, delay and amount of choices has been investigated. In neither of these cases a significant correlation was found, which indicates that the start time has not influenced the choices participants made or the outcomes of their trips in terms of travel time or delay.

8.5.2 Details of travel time and delay

The travel time is shown as boxplot in figure 8.3. The line in the middle of the box represents the median and the box contains the middle 50% of the data. The range of these middle two quartiles is known as the inter quartile range or IQR. The whiskers at both ends extend to 1.5 times the IQR, with the dots indicating outliers in the data. For scenario 2 both incidents cause a larger variation for trips without information, because the difference between the 25% and the 75% quartiles is much larger compared to the trips without incidents. This also applies to incident 3 and 6. Incidents 4 and 5, however, show only a small difference. This, again, is most likely caused by the incident not being on the preferred route, which meant participants did not have to select a new route, even though they had no explicit information on the incidents.

When providing information the travel times show a reduction of standard deviations. It is clear that the colored areas between the 25% and 75% quartiles is much smaller than in the case of participants having no information. In general, there even is another improvement between textual information and dynamic route guidance, where the latter has a lower median.

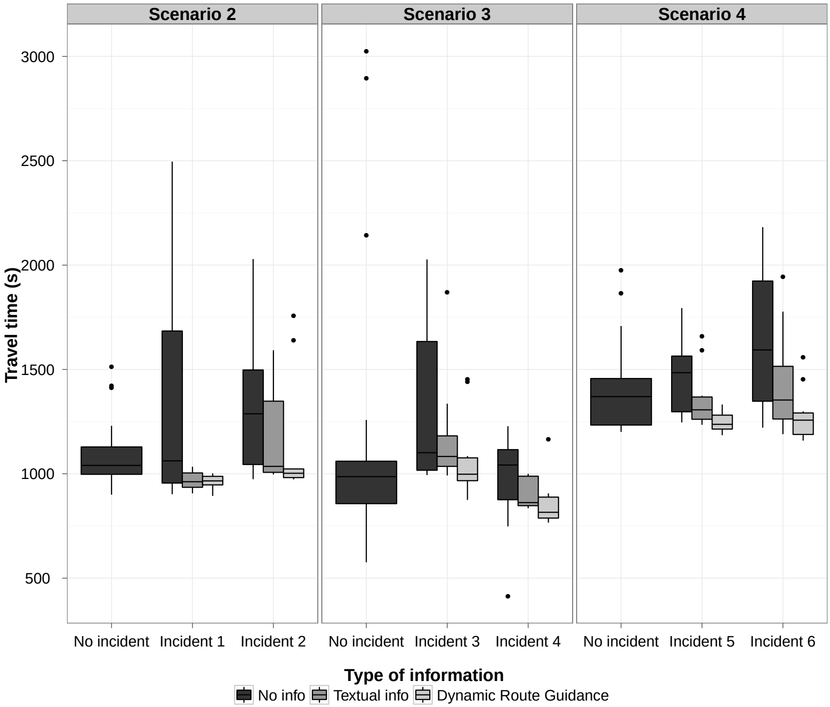


Figure 8.3: *Boxplot of travel times (s) per scenario, type of guidance and accident situation*

When looking at the boxplot for delay (figure 8.4), scenario 2 differs from scenarios 3 and 4. Scenario 2 has a large increase in variation in the delay for the situations with incidents and a large increase of the median for incident 2 without information, whereas this is not the case for scenarios 3 and 4. This indicates that for scenario 2, participants still used the same route as without incidents, whereas for the other situations, participant adapted to the specific circumstances. The fact that the freeflow travel time for scenario 2 is almost the same, is also proof for this.

Interestingly, the boxplot for delay shows that there is not much difference between the various types of information, apart from the situations in scenario 2. Scenario 3 and 4 both show only small differences in median for having information or not. An improvement in travel time thus is not caused by having a lower delay, but by having a lower freeflow travel time, which is in agreement with the results in table 8.8. Scenario 2 is different in this case, where especially for incident 2 and dynamic route guidance a large difference is found in terms of variation. The average and median are not that different for the overall travel time, but the freeflow travel time is. This means that participants selected a longer route, but this route has less delay.

To show the effects of type of traffic information on the travel time or delay for the trips with incidents Analysis of variance (Anova) is used. Because of the experimental setup with multiple trips for each participant with the combination of traffic information and incident it resulted in a repeated-measures Anova. For

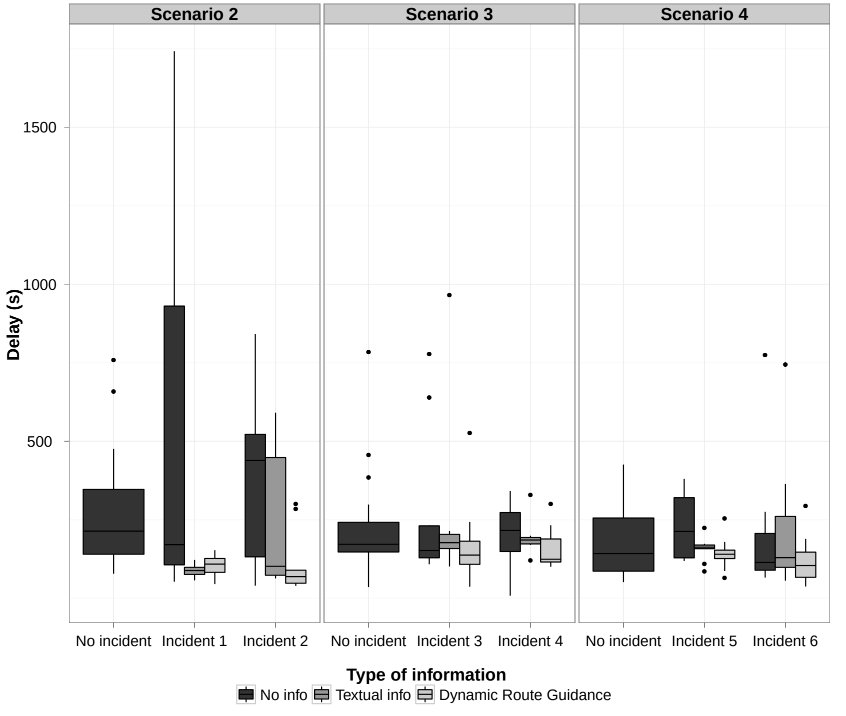


Figure 8.4: *Boxplot of delay (s) per scenario, type of guidance and accident situation*

the various incident situations, the outliers have been removed to represent the right measurements. The trips where the right destination was not reached, or where a participant obviously (and noted during the experiment) made a wrong choice, have been removed from the dataset. This resulted in an unbalanced dataset for the various incidents. The type of information is used as a within subjects factor. No between subject factors have been used, since these variables (such as gender or amount of kilometers per year) were not available for all trips (because of the removal of the outliers), or in too low number to calculate valid statistics. Therefore it was a one-way Anova. For these variables a Shapiro-Wilks test and Levene's test were performed to respectively test for normality and equality of variances. Although the results of the Shapiro-Wilks test indicate that the results are in most cases not normally distributed, it was chosen to continue with the Anova's, as this is the least critical assumption and a non-parametric test cannot be used for unbalanced datasets. Levene's test for homogeneity of variances showed a significant difference for incidents 2 and 5. This means that for the three types of information, the variance is not equal in these cases. A non-parametric variant of the Anova would be more appropriate in these cases, but this test (Friedman's Test) is not capable of repeated measures experiments. Therefore an Anova for these two situations has been conducted. The results of the Anova are given using the calculated value for F with the degrees of freedom for the treatment combination and the degrees of freedom for the errors as $F_{\text{df treatment,df error}}$. This value is checked against the critical values of the F-distribution.

When looking at the effects of traffic information of the various incident situations, a significant effect has been found for incidents 2 ($F_{2,23}=5.3$, $p<0.05$), 4 ($F_{2,24}=11.4$, $p<0.01$) and 5 ($F_{2,23}=8.5$, $p<0.01$). The boxplot for travel time with these incident situations shows that indeed in these situations traffic information had a substantial effect. The largest differences between the incidents show a significant different mean when looking at the effects of type of traffic information. For delay the same significant effects of traffic information have been found for incident 2 ($F_{2,23}=5.5$, $p<0.05$), 4 ($F_{2,24}=14.1$, $p<0.001$) and 5 ($F_{2,23}=6.7$, $p<0.05$).

The situations where traffic information was available (either as textual information or dynamic route guidance) are showing clearly that being able to use information has an effect on average trip lengths, travel times, delays and their respective standard deviations. In all scenarios, the average travel time for trips with textual information and dynamic route guidance is lower compared to the situations where no information was available, but had incidents. For example, for incident 5, the difference between having textual information and no information is almost two minutes. Dynamic route guidance again improves this average travel time by almost two minutes. In total, a participant could reduce his travel time by almost four minutes, on a trip of 25 minutes, which is around 16% improvement in travel time. For incident 6, this improvement is even larger (over 5 minutes). The experienced delay is also reduced for both incidents, but the main improvement is caused by participants selecting routes with a lower freeflow travel time, thus a faster route.

Traffic information thus not only allows participants to select routes that go around incidents, they are also more rapidly improving their routes in terms of (freeflow) travel time. Since in some incident situations the average travel time is lower than in the situations without incidents, there is some learning effect of the participants, also induced by the traffic information. Overall, not just the averages are reduced. Also the standard deviations for the travel time, freeflow travel time and delays are reduced by having traffic information compared to having no information, for the situations with incidents. This decrease indicates participants selecting more of the same routes.

The average delay has also been calculated as percentage of the average total travel time which can be found in figure 8.5. For situations without information and incidents, the delay is between 25 and 13% of the total travel time.

Apart from incident 6, an incident increases the delay in the travel time, where especially scenario 2 shows a large increase for incident 1. This is most likely due to the fact that the incident has a large influence on the travel time on the highway route which 4 participants used. For this incident, it is clear that information caused all the participants to avoid the highway, thus greatly reducing the percentage of delay in their travel times. Incident 2 shows an increase of participants selecting the fastest route, which avoided the incident, therefore reducing the delay. Incidents 3 and 4 show only a small increase in the delay as percentage of the total travel time when no information was available. Traffic information hardly improves this. These small differences are partly due to the mistakes participants made because of the north-south issue. Scenario 4 also only show small differences in the percentage delay has in the total travel time. Incident 6 in this case shows curious results, as participants experience more delay in their total travel time for textual information.

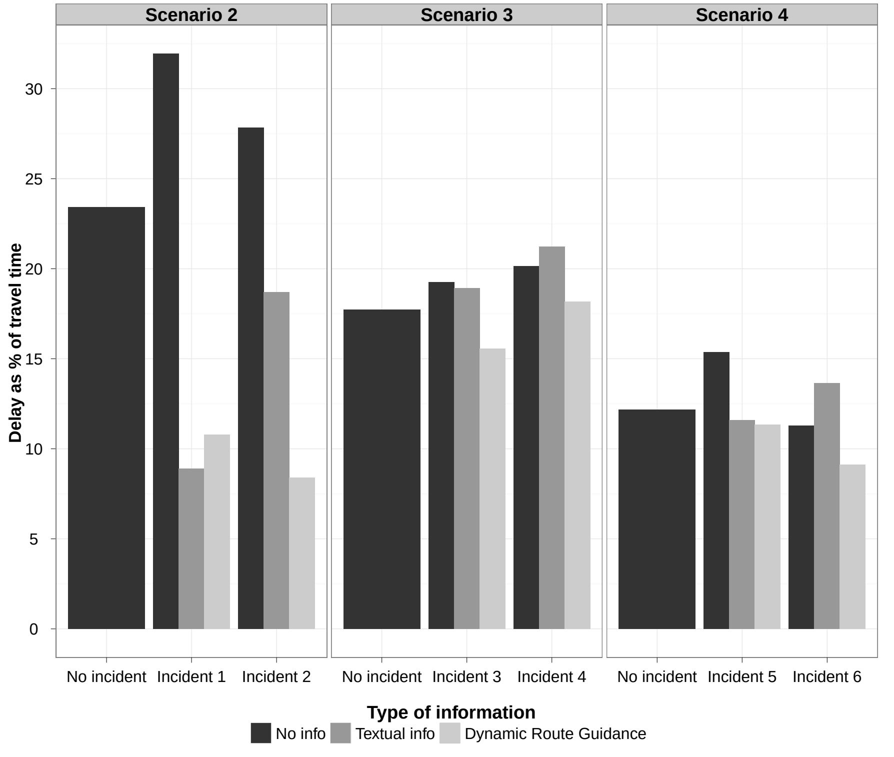


Figure 8.5: *Delay as percentage of the travel time*

8.6 Discussion

The AIDA RCS was tailored for the experiment described in chapter 7 to investigating the effects of traffic information on route choice when accidents occur in the urban road network. To this end, the route choice simulator has proven to be a valuable method for gathering the necessary data. The architecture used, based on the microscopic traffic simulation model Paramics provided a detailed simulation environment, which is capable of creating realistic traffic environments. Together with the API of Paramics, it is possible to extract all the necessary data of the traffic simulation and ask pre and post trip questions to the participants. Paramics also served as a valuable environment to build the road network and traffic situations.

In this experiment, a hypothetical network was used. The choice locations were depicted using screenshots of these junctions from Paramics, in which the junction could be seen. As no surrounding buildings were used, it was not possible to navigate in the road network using striking buildings for participants. The information available was the maps they were given beforehand and the information in the AIDA RCS client itself. The maps are expected to offer enough information for participants to perform their wayfinding in the network. However, the experiment did not focus on wayfinding during non-recurrent traffic situation, but targeted the effects of these situations combined with traffic information on route choice behaviour. Based on this goal, we feel that the lack of realistic screenshots (including surroundings) does not mean the results are invalid.

As mentioned, in scenario 3, the orientation of the routes was from north to south. This caused some participants to have difficulties distinguishing left and right, resulting in “wrong” or “weird” choices. In these situations, the AIDA RCS client presented contrasting information. The map of the total road network did not rotate with the direction of the participant, but always had north at the top. The map of the upcoming junction did rotate with the direction of the participant, which thus shows opposite of the other map if the participant is driving south. The results of scenario 3 are thus different compared to the other scenarios. However, the number of routes used in total and per participant, and the averages of these routes with respect to travel time, delay, length are equal to the other two scenarios. This indicates that for an aggregated level, the difficulty of the orientation has not caused significant differences.

The number of traffic messages a participant received were rather high, especially in scenario 4. This is partly due to the high update frequency of one minute, but is also due to the traffic jam detection, as mentioned in section 6.3.3. It is very well possible that a traffic jam which was visible as one traffic jam was detected as two separate traffic jams. The minimum length for a traffic jam to be reported partly resolved this issue, because a traffic jam had to be at least 1000 m in length. Short links which lead to a traffic jam but were not selected as part of it, would be ignored for creating the traffic messages. Based on this workaround, and the possibility for a participant to combine the messages with the overview of the road network, it is expected that there has been no influence on the routes selected by the participants.

The three scenarios were different with respect to their origin and destination and the locations of the accidents. With respect to the other elements (type of guidance, random seed and incident or not), they were all equal, and used the same order of these elements for all participants. In order to prevent learning behaviour by using traffic information and/or route guidance, together with the knowledge of incidents and their locations, the order was not randomized between participants. For instance, if a participant started with a trip with route guidance and traffic information, he/she would be informed of the fastest route automatically, together with information of a possible incident. This information is very likely to be used in consecutive trips, which makes it impossible to retrieve the participants behaviour if no guidance or traffic information is provided. The order that is used in the scenarios is now providing more information to the participants, which is confounded with the learning because of previous trips and experiences in these trips. This means the results for traffic information cannot be completely distinguished from a possible learning effect, when looking at aggregate statistics of complete trips. Using individual choices, where the number of the trip in the sequence of all trips is one of the explanatory variables, will show if there is some learning (see chapter 9).

In total three scenarios are used for data analysis, each with two incidents. For each scenario, at least two reasonable routes were determined, on which these accidents were located. If a participant has a different perception of the road network and selects other routes, it is likely he or she will not directly experience the incident, or indirectly its consequences for surrounding roads. To minimize this possibility, incidents were placed on the two most reasonable routes, so a participant was very likely to run into such an accident at least once. As the overview of routes showed, this indeed happened, as just a few trips did never run into an incident. However, it also means that using this combination of incident locations, it is not possible to see if participants are influenced by incidents that are not anywhere near their reasonable routes.

The sequence of trips made by a participant starts with two trips in which neither information nor accidents are used. These two trips were meant to allow for a participant to find a certain preference for a route. It is expected that two trips are not enough to enable a participant to fully determine his or her preferred route, which is more likely to need around 10 trips. To acquire such a preferred route, the experiment would last more than twice the current time. This would mean a big increase on respondent fatigue, which is undesirable. This also means that incidents could not be located on a preferred or habitual route, which would allow for the estimation of route switching behaviour. The data that is gathered with this experimental setup allows for a possible rerouting if the participant selected on of the reasonable routes on which the accident is located, but not if the other route (or any other route) was chosen. As the goal of the research is not to investigate such route switching, it was chosen not to relate the routes selected to the reasonable routes.

The experiment focused on the possible effects of traffic information and/or route guidance and incidents. This means that many other items of interest of ATIS are not studied, such as reliability and timeliness of traffic information. Neither are multiple algorithms implemented in the route guidance, apart from the fastest route. It is assumed that participants are interested in the fastest route, as the scenario concerns a morning commute, which was also found as the most important type of information in chapter 5. For route choice behaviour, the interest is with the effects of both ATIS and incidents on the total routes and individual choices. Although behavioural aspects such as risk aversion can be of influence on the choices and routes, it would increase both the duration of the experiment and the number of participants necessary for valid results, which was not possible for this research.

For the route choice study, an invitation was send to friends and colleagues. The sample is not based on a sampling strategy, and the participants are not a representative selection of the Dutch population of car drivers. The results cannot be interpreted as a true representation of what would be found in a real life situation. Together with the small sample size, it is not possible to achieve statistically significant effects of gender or age on routes selected and the characteristics of these routes. The results are only given for the whole group of participants, regardless of the socio-economic characteristics of the participants. The consequence of this is that the results as such cannot be transferred directly to real life situations. The behavioural implications and the overall size and direction of effects are however possible to transfer. This is based on the assumption that the AIDA RCS provides a realistic environment in which the participant would behave as they would in real life. The post-trip questions that each participant answered once arrived at the destination are a proof of this validity (face-validity). Together with the strong effects found for both incidents and traffic information in each scenario (internal validity), also indicate that the AIDA RCS is a valid tool, which provides results which can be transferred to real life.

The results show a significant reduction of travel time when participants are able to use a form of ATIS. As no habitual behaviour is estimated, the reduction of travel time cannot be related to the travel time which a participant otherwise would have experienced. As such the direct influence of a possible rerouting is not described, or the probability a participant will reroute based on the combination of incidents and/or traffic information. When looking at the overview presented in Toledo & Beinhaker (2006), the range of travel time reduction found in this study is in line with the findings of other studies. In Bogers (2009) a reduction of average travel time for a highway network with three alternatives was found of around 8%, when participants received both en-route and ex-post traffic information. However, no accidents are used in this case. Compared to the reductions found in this research, this reduction is smaller.

In a study which did include incidents (Emmerink et al., 1995), there is a reduction of maximally 20% (for a market penetration of 50%). However, this result is achieved by simulation in which various assumptions are used in the models route choice and information usage. It differs from this study, as here the travel times are experienced by the participants. In a more recent study using a simulated environment Abdalla & Abdel-Aty (2006), the amount of travel time savings range from 20.7% to 44.7% for pre-trip information without advice and en-route traffic information with advice respectively. In this situation no incidents are used, in stead different weather conditions are influencing the traffic situation. As the participants were familiar with the road network, the amount of travel time reduction is rather large, compared to the reductions found in this study.

8.7 Summary

This chapter focused on the general results of the experiment undertaken with the AIDA RCS. This provides an answer to the research question from section 1.2.3 on the effects of traffic information on route choice during non-recurrent traffic situations. As explained in chapters 6 and 7, only accidents are used for the non-recurrent traffic situations, and they occur during a morning rush hour.

First an overview of the participants in the experiment was given. The participants are not a proper representation of the "average" Dutch car driver, as it differs for all variables. Especially age, gender and level of education are different, which is caused by many direct colleagues taking part in the experiment. However, it is not expected that this caused significant differences in the choices made during the trips.

To investigate the face-validity of the AIDA RCS, each participant was asked a few questions after finishing the trip. These questions concern the general experience and whether the participant would make the same choice in a real life situation. Both of these questions indicate a high level of validity of the AIDA RCS and the experiment. Most participants said the trips to be either realistic of very realistic, and most also would make the same route choice in real life situations. Just scenario 3 with a north south orientation had a lower level of validity, which is easily explained as participants had trouble to navigate with this orientation.

A graphical overview of the routes, showed that the participants selected a different route when first receiving traffic information, and further seem to “improve” on their route choice when driving with dynamic route guidance. The number of routes used and the frequency hereof is also reduced with extra traffic information. Traffic information and route guidance both cause a change in route choice, as participants show less variation in their choices. This indicates that a form of concentration occurs. The amount of routes used varied per scenario and per incident, as in each scenario one of the incidents was on the route that most participants seemed to prefer. These accidents caused participants to reroute and seek a way around the incident or the congestion caused by it. Scenario 2 had the most obvious routes, as only 8 routes were used. Scenario 3 and 4 were different in that respect, as in scenario 3 a total of 26 routes was used and 29 routes in scenario 4. This is most likely due to the location of the origin and destination in scenario 3 and 4, and also to the orientation in scenario 3.

Incidents caused an increase of 9.7% in the experienced travel times in general when no information was provided. The amount of delay increased by 30%. As participants were not informed of the accidents, they were not able to anticipate the possible delays, and often experienced the congestion of the accidents or the accidents themselves, hence the increase in travel times and delays. The freeflow travel time did not show a significant increase, nor did the length of the routes used.

Traffic information allows participants to reduce their average travel time in incident situations by more than 13%, and dynamic route guidance further reduces their travel time with 6%. This is a reduction of over 4 minutes on a travel time of 22 minutes. Depending on the specific incident situation, dynamic route guidance can reduce the travel time with 32%. Without any information, there is an average delay of around 5 minutes, regardless of the scenario or the specific incident. Traffic information reduces the delay with 40% and dynamic route guidance shows a total decrease of 58%. The largest contribution (around 70%) to the reduction in travel time is thus caused by a reduction of the amount of delay, because participants avoid the incident or the congestion caused by it. The travel time in the trips made with dynamic route guidance consists for around 10% of delay. In case of three incidents, which all are on the preferred route of participants, both types of traffic information have a significant effect on both travel time and delay. Traffic information and dynamic route guidance also reduce the variation of travel time, delay and length of routes, which confirms the concentration effect found in the graphical overview of the used routes.

You must keep your mind on the objective,
not on the obstacle.

— William Randolph Hearst

Chapter 9

Route choice experiment - Choices and choice models

9.1 Introduction

This chapter focusses on the individual choices of the participants in the experiments with the AIDA RCS and thus concerns a more detailed level of understanding the effects of incidents and/or traffic information. At each junction in the road network, the choice was recorded, together with the possible alternatives, and, if provided, the advice provided by the dynamic route guidance. This enables us to investigate a possible relation between the choices and the circumstances of the trip.

The individual choice is a choice between a discrete number of alternatives or exits at a junction. Each exit has at least one route which leads towards the destination. Using discrete choice models, it is possible to relate the choice made to the attributes assigned to an exit. The theory behind these models was described in chapter 4 with paying explicit attention to discrete choice models in a route choice context (see section 4.3). The choice models will provide an answer to subresearch question 3c: *Which attributes are important in route choice behaviour?* and the main research question: *What are the effects of traffic information on route choice during non-recurrent traffic situations?*

In section 9.2 the individual choices are discussed and related to the attributes of the exit chosen. The choices are also discussed for the various incidents, types of traffic information and scenario. The choice models are presented in section 9.3. This section starts with the response to one of the post trip questions regarding the most important factor in the route choice of that trip (section 9.3.1), follow by a description of the datasets used and the approach for the development of the models in section 9.3.2, followed by the development of a base model, based on a linear utility function for the alternatives (section 9.3.3). It applies this utility function to a complete dataset and a dataset in which only the strategic junctions are used. The base models are further extended with two specific correction factors for the amount of overlap of the alternatives: the Path Size and Path Size Correction factor. This section also introduces three variables for the inclusion of traffic messages, incidents or advice to the base models. Next is a discussion of the internal validity of the estimated models (section 9.3.4), by comparing the predicted and observed probabilities for choosing a certain exit. The last part of this section concerns relaxing the assumptions of a multinomial logit model (MNL), by accounting for heterogeneity and heteroscedasticity (section 9.3.5). The results will be discussed in section 9.4, followed by a summary in section 9.5.

9.2 Choices

During each trip, a choice had to be made at each junction. Three types of choices were possible:

1. no choice possible;
2. participant choice;
3. no participant choice.

In case no choice was possible, it meant that a link in the RCS was connected to only one other link, essentially giving no choice. It was also possible for participants to make no choice. If the participant did not make a choice (in time), Paramics made the choice for the participant. This choice was automatically the fastest route, which was mentioned by some of the participants, but never reported as problematic. Of course, if a participant made a choice, this was also recorded. An overview of the amount of different choices is given in table 9.1. In total 2.945 choices have been recorded excluding the choices in the warm-up scenario and the non participant choices. 42% of these choices were made in scenario 4, 27% in scenario 3 and 31% in scenario 2.

In this case, the results for the warm up scenario have explicitly been taken into account. Comparing the amount of "no participant choice" for this warm up scenario with the other three scenarios, it is obvious that the warm up scenario has a much higher number of missing choices. This is most likely due to a learning effect. During this first trip, participants had to get accustomed to the AIDA RCS and the speed with which they were expected to make choices. It is therefore likely that many participants would miss one or more choice moments. In the consecutive trips, this is less likely to happen. However, as previously discussed, scenario 3 meant an extra complication for a lot of participants, because the trip was in a different direction then the map in the RCS client, which caused them to have more trouble choosing. This is also visible in a higher percentage of missed choices. Interestingly, this percentage decreases when more information becomes available. In the case of dynamic route guidance, it is reasonable to expect that in case a participant was not able to choose by himself on time, he would select the advised

Table 9.1: *User and missed choices per scenario and type of guidance*

Type of guidance	**Participant choice**	**Warm up**		**Scenario 2**		**Scenario 3**		**Scenario 4**	
No information	No			22	5.1%	50	11.6%	18	2.8%
	Yes			414	94.9%	382	88.4%	629	97.2%
Textual information	No			5	2.0%	17	7.4%	8	2.6%
	Yes			240	98.0%	212	92.6%	304	97.4%
Dynamic route guidance	No	46	11.8%	1	0.4%	13	5.7%	4	1.3%
	Yes	344	88.2%	249	99.6%	215	94.3%	300	98.7%

exit. In the case of textual information, the reduction in missed choices is more likely to be caused by a learning effect, as participants learned how to navigate ”upside down”. Aggregated over all types of information and scenarios, an average of 3.56% missed choices were recorded.

For the choices where a participant made his or her own choices, a more in depth analysis has been performed. First, the choices made by participants have been compared with the exit of the fastest route at that time. A participant could either choose the fastest route or not. The results are give in table 9.2.

In general, when looking at the various scenarios, it is clear that the more advanced the information is, the more a participant is likely to select the exit at a junction which matches with the fastest route at that time. Without information, participants selected the fastest exit in 84.3% of the choices. For textual information, where information on the incidents and traffic jams was provided, this increases to 87.2%, whereas for the dynamic route guidance, where an exit was advised, based on the current fastest route, 94.2% of the choices matched the fastest route.

Scenario 4 shows a slightly different pattern between no information and just textual information compared to the other scenarios. In these cases, the percentage of participants that did not choose for the fastest exit is more or less equal. This might be due to the many possibilities for this scenario, especially at the “Koningshuis” area. The difference in travel times in this area is very small, so it is likely that participant did not select the fastest of the routes available here. The increase for the same scenario when looking at the Dynamic Route Guidance, shows that participants were much more likely to follow the advice given, and selected the fastest routes, even in the “Koningshuis” area. For scenario 3, the increase for choices on the fastest route is smaller. This can be explained with the “wrong” choices participants made, because of the north-south orientation. However, even with these problems, participants are quite capable of selecting the fastest exit. For scenario 2, this is even more obvious. For the trips made with Dynamic Route Guidance in this scenario, as little as 2% of all choices did not select the fastest exit. It means that participants trusted the route guidance to a very high level, as the route guidance always provided advice for the fastest exit at the moment of choice.

The results for choices aggregated per incident (table 9.3) clearly show an influence of incidents on the choices for the fastest exit. Incidents 2 and 3 have a much lower percentage of choices for the fastest exit, regardless of type of information. The location of these incidents thus greatly influence the behavior of participants. For

Table 9.2: *User choices for the fastest exit per scenario and type of guidance*

Type of guidance	Fastest exit	Scenario 2		Scenario 3		Scenario 4	
No information	No	52	12.6%	72	18.9%	101	16.1%
	Yes	362	87.4%	310	81.1%	528	83.9%
Textual information	No	19	7.9%	29	13.7%	48	15.8%
	Yes	221	92.1%	183	86.3%	256	84.2%
Dynamic route guidance	No	5	2.0%	26	12.1%	13	4.3%
	Yes	244	98.0%	189	87.9%	287	95.7%

Table 9.3: *User choices for an exit matching the fastest route per incident situation*

Scenario	Incident situation	Amount	Percentage
2	No incident	176	90,2%
	Incident 1	315	96,0%
	Incident 2	336	88,4%
3	No incident	159	82.4%
	Incident 3	258	79,8%
	Incident 4	265	91,4%
4	No incident	272	86.3%
	Incident 5	396	88,0%
	Incident 6	403	86,1%

incidents 1 and 4, participants were more able to select the fastest exit, compared to the situations without incidents (for all scenarios). For incidents 5 and 6, this difference is only small.

For trips in scenario 2 and 3 without information one of the incidents (either 2 or 3) is likely on the preferred route (as discussed in section 8.4), because the percentage of choices for the fastest exit is lower than the trips without incidents (see table 9.4). The percentages for the fastest exit in scenario 2 with no information are significantly different ($p<0.05$) between the incidents, based on the Kruskal Wallis Rank Sum test. Scenario 3 without information shows no significant differences for the two accident situations. Interestingly, the percentage for the fastest exit with the other incidents (1, 4, 5 and 6) in all scenarios is not significantly ($p>0.05$) different from the trips without incidents. This indicates just a small learning effect for trips made without information, as the trips with incidents are made after the trips without incidents.

For each incident situation (figure 9.1), the differences between the percentages of choices for the fastest exit between the three types of information have been investigated. Incidents 2, 4, 5 and 6 show a significant difference ($p<0.05$), which means that traffic information and dynamic route guidance had a significant effect on the propensity of the participants to choose the fastest exit. When comparing the individual combinations for the incident situations with a significant difference, the differences are largest between the trips without information and the trips with dynamic route guidance.

A further analysis can be made of the choices, for the trips where dynamic route guidance was available. During these trips, the participants received advise on which exit to take, which was based on the fastest route. For the choices made, a comparison was made for the chosen exit and the advised exit (table 9.5). Of the incidents, those scenario 3 show the smallest percentage of choices where participant selected the advised route. Again, this is most likely due to the problems participants had with the orientation, in particular for incident 3. For the other incidents the percentage of choices of participants matching the advice is very high, close to 100%. Combined with the other results and the self-reported validity of the experiment, it is likely participants trusted the advice given.

Table 9.4: *Percentage of choices for the fastest exit.*

Scenario	Type of guidance	Accident situation	Percentage
2	No info	No incident	91.71%
		Incident 1	91.04%
		Incident 2	77.70%
	Textual info	Incident 1	96.92%
		Incident 2	86.73%
	Dynamic Route Guidance	Incident 1	99.23%
		Incident 2	97.57%
3	No info	No incident	82.31%
		Incident 3	78.00%
		Incident 4	83.14%
	Textual info	Incident 3	80.90%
		Incident 4	94.37%
	Dynamic Route Guidance	Incident 3	84.56%
		Incident 4	94.40%
4	No info	No incident	86.23%
		Incident 5	83.16%
		Incident 6	79.43%
	Textual info	Incident 5	84.81%
		Incident 6	83.69%
	Dynamic Route Guidance	Incident 5	96.00%
		Incident 6	95.53%

Table 9.5: *Amount of user choices for a next link matching the advised link (with dynamic route guidance) per incident situation*

Incident situation	**Amount**	**Percentage of all choices**
Incident 1	110	99,1%
Incident 2	134	97,1%
Incident 3	98	83,1%
Incident 4	91	93,8%
Incident 5	142	95,9%
Incident 6	145	95,4%

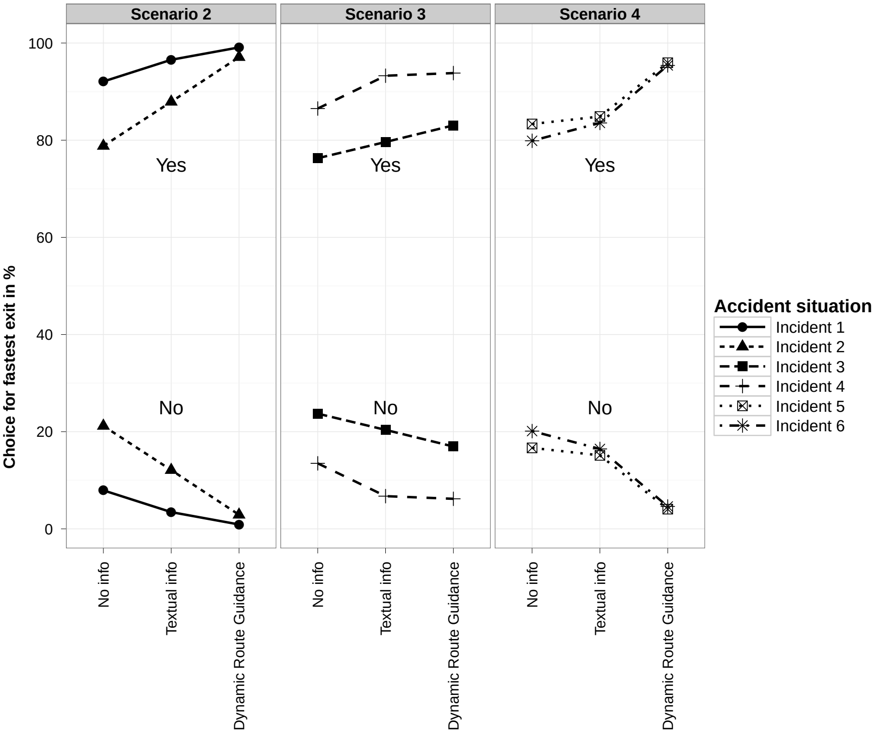


Figure 9.1: *Percentage of choices for the exit on fastest route, per scenario, accident situation and type of guidance (for choices made by the participant)*

9.3 Choice models and attributes

This section concerns the individual choices of the participants in the route choice experiment. Data of each choice at a junction is used in estimating the proposed route choice model (see section 4.3.4. The section starts with a presentation of the attributes for route choice that were reported by the participants themselves (section 9.3.1). These were a starting point for the utility function that is used to describe the three alternatives in the exit choice model. In the approach, the data, the utility functions and the available attributes are described (section 9.3.2). This section also describes the assumptions used in the utility models that were applied in the estimation of the models. The first step in the approach is to find and estimate the base utility functions (section 9.3.3). The second step is to expand the basic utility function with additional attributes (section 9.3.4) for which two different sets are used. The first relates to one of the assumptions of the multinomial logit model (i.i.d.), which is violated in routes that have overlap. Two correction factors have been applied and tested for using the base model. The second set of attributes concerns attributes for accidents, traffic information and route guidance. The resulting utility functions have been used to predict choices for a part of the dataset not used in the estimation, which is a measure of the internal validity of the estimated model (section 9.3.5). The last part presents the results of the estimation with other approaches in the modelling, which relax the assumptions made in the MNL model (section 9.3.6).

9.3.1 Self-reported attributes

One of the questions asked to each participant after a single trip is the question on the most important factor in the route choice of that trip. As such it indicates the reasons behind the route choices made. The results are given in table 9.6, where a distinction is made between the three types of guidance a participant received during each trip.

As expected, the type of traffic information is of significant influence on the most important factor in route choice (Chi-square test shows $p<0.05$). This also applied to the different incident situations, where especially the trips without incidents (and thus without information) are of significant influence. The scenarios did not have a significant impact on the selection of the most important factor.

Overall, traffic information or route guidance is the most important factors, which increases to 63% of the responses when dynamic route guidance was provided. This indicates that participants have a high level of trust of the traffic information and/or route guidance. The information must have been interpreted as reliable by the participants, otherwise it would not have been the most important factor in their route choice. In the case of no information, there is a spread over all options, where previous experiences, allowed speed and distance are selected most often. Interestingly, an accident on the network was selected only 5.9% of the trips as the most important factor (especially for trips with incident 3 and 6). This could indicate that participants are quite indifferent to accidents on the network. However, because no definitions have been provided to the participants, it could

Table 9.6: *Answers for the question "What was the most important factor in your route choice?" per type of guidance*

	Answer							
Type of guidance	**Accidents on the network**	**Traffic information or route guidance**	**Previous experiences**	**Allowed speed on this route**	**Distance of this route**	**Travel time of this route**	**I don't know**	**Missing**
No information	4.2%	8.3%	21.7%	22.5%	22.5%	13.3%	4.2%	2.5%
Textual information	10.0%	45.0%	16.7%	0.0%	13.3%	13.3%	1.7%	0.0%
Dynamic route guidance	5.0%	63.3%	16.7%	1.7%	0.0%	10.0%	0.0%	3.3%
Total	5.9%	31.4%	19.2%	11.7%	14.6%	12.6%	2.5%	2.1%

also mean that most participants chose for traffic information instead of accidents, because the accident was made known to them by means of traffic information. Adding to that, accidents are chosen also as a factor in trips where no information was available, but most often in trips with traffic information available.

The importance of traffic information and route guidance was also found for the compliance rates found in section 9.2. The participants also mentioned this when specifically asked. It is expected that having an advice or traffic information for a route using that exit will be a very important attribute of the routes in the choice models.

9.3.2 Datasets, approach and linearity

Datasets

Using the individual choices at each junction it is possible to estimate the exit choice model proposed in section 4.3.4. Such a model will estimate the probability of choosing one of the alternatives available at that location. In this application, the alternatives are the exits at a junction. A U-turn is not available. The road network was designed to use no more than four links for a node, which makes a maximum of three alternatives.

These alternatives are labelled as:

- left;
- straight;
- right.

As not all junctions have right-angled exits, the exits which are closest to the right-angled values, are set to the respective label. The exit with an angle closest to 180° is labelled as the straight exit, etc.

Table 9.7: *Availability and choices for each dataset*

	Complete dataset		**Strategic dataset**	
	Available	Chosen	Available	Chosen
Right	1850	521	820	324
Straight	2252	1580	815	415
Left	2010	416	605	166

The choice set thus consists of a maximum of three alternatives (in a T-junction, one of the alternatives was set to unavailable), for which it is necessary to describe the characteristics or attributes. It is expected that drivers select an exit based on the characteristics of the route that starts with that exit. For each exit, the fastest route towards the destination was calculated, using the A^* algorithm based on the actual travel times, as it is expected that for the morning commute scenario, all participants are interested in the fastest route (based on the findings in chapter 5). For each of the routes that are calculated, the following attributes are defined: (freeflow) travel time, delay, length, amount of junctions, number of left turns and existence of incidents, traffic information or advice. These attributes are to be used in the utility function of the choice model.

The resulting choice model is a model that predicts the probability for an exit, not for a route for the (remaining part of the) trip. At each junction, the same approach is used to predict the probability of choosing an exit, based on the attributes of the routes that start at that exit.

Two datasets are used for estimation. The first consists of all choices at junctions with two or more exits (note that by design, no junctions exist with more than four arms) that provide a route that reaches the destination. Some junctions with three exits have an exit that is a dead-end street, which renders that exit unavailable for the choice set at that junction. The second dataset consist of a selection of all the choices, called the strategic dataset. The selection is made using the predetermined possible routes for a scenario (see figures 7.5b, 7.6b and 7.7b). This lead to a total of 8 junctions from scenario 2, 9 junctions from scenario 3 and 11 junctions from scenario 4. For the complete dataset it means that 2517 observations are available, whereas for the strategic dataset this reduces to 905 observations (see table 9.7).

Utility functions

As there is a huge number of utility functions possible to be estimated from the dataset and the attributes, a step-by-step approach is used. The first step is to start with a number of basic attributes, which are often used in route choice models. These attributes are analysed with respect to their values and distribution, if it concerns continuous values such as travel time or delay. If deemed applicable, these attributes are selected for a transformation, to create a linear set of values. This model is further applied to both datasets, with and without extra attributes. The first concerns attributes of the routes, followed by the correction factors described in section 4.3.3. The final addition concerns the attributes for incidents, traffic message and advice. The final set of attributes is then applied to different datasets for estimation and prediction, as an indicator for the internal validity of the model.

The same base model is also used to account for heterogeneity and heteroscedasticity in the dataset.

The following basic linear-in-parameters utility function is used to determine the most appropriate attributes of the alternatives.

$$V_i = ASC_i + \beta_{xi} X_i \tag{9.1}$$

in which:

$$\begin{aligned} ASC_i =& \text{Alternative specific constant for alternative i} \\ \beta_{xi} =& \text{Vector of utility coefficients for attribute x and alternative i} \\ X_i =& \text{Vector of attributes for alternative i} \end{aligned}$$

For this function the vector of parameters X_i consist of variations of (based on Ben-Akiva & Bierlaire (2003)):

TT_i Travel time from the current junction to the destination, including delay;

$FFTT_i$ Freeflow travel time from the current junction to the destination using allowed speed and length;

D_i Delay from the current junction to the destination (travel time − freeflow travel time;

JC_i Amount of junctions between the current junction and the destination (equals the number of choices to make);

LT_i Amount of left turns between the current junction and the destination;

C_i Dummy describing if the alternative decreases the crow fly distance to the destination;

S_i Average speed of the alternative, using the speed limit, averaged over the length.

Non-linearity

For the continuous attributes (TT_i, $FFTT_i$ and D_i) a transformation can be applied, as these attributes are not necessarily linear in itself as is assumed for the MNL model. Non-linearity means that the value of an attribute is not on a straight line but has a non-linear effect on the dependent variable (utility in this case) as a the effect of an additional unit of travel time for a low value of travel time is different from a high value of travel time. Such a non-linearity can be overcome using a transformation. Most known examples are the log and Box-Cox transformation (Box & Cox, 1964). The Box-Cox transformation is given in equation (9.2).

$$X_i(\lambda) = \begin{cases} \dfrac{{X_i}^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0; \\ \log X_i & \text{if } \lambda = 0. \end{cases} \tag{9.2}$$

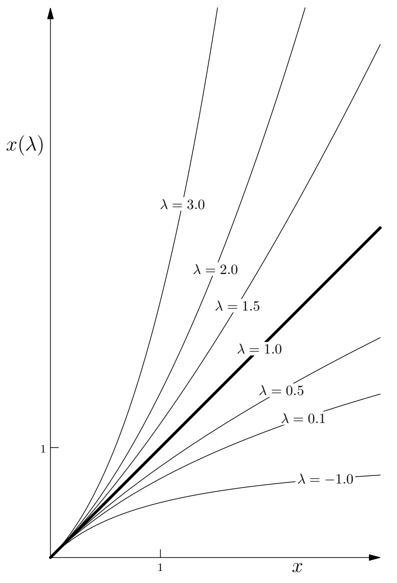


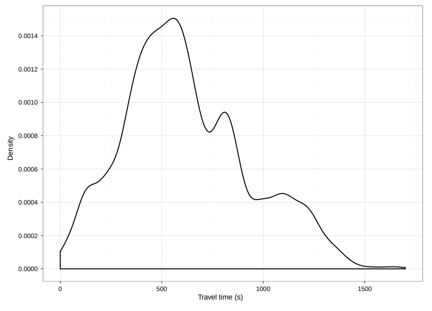
Figure 9.2: *Box-Cox transformations for x with several values for lambda*

The parameter λ is a measure of influence of additional effects. If λ is below 1, it means that the effect of an addition unit of for example travel time at a higher level of this travel time is smaller than at lower levels. The other way around, a λ of over 1 means that the effect of an additional unit of travel time will be higher at higher levels of this travel time. The Box-Cox transformation is often applied to reduce anomalies in a variable, such as non-additivity, non-normality or heteroscedasticity. In the case of travel time, it is clear that it cannot be normally distributed, because it is strictly positive. This is confirmed by the Shapiro-Wilk test for normality ($p<0.05$) which makes it a proper candidate for a transformation. The Box-Cox transformation has been used in discrete choice studies before and is easily estimated using BIOGEME (Bierlaire, 2008).

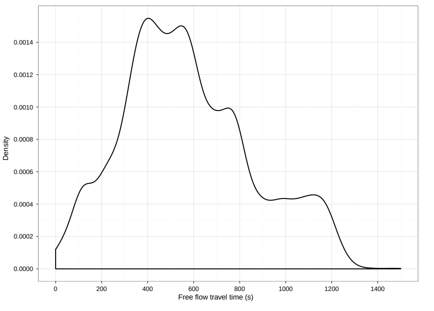
The Box-Cox transformation is a generalization of various power transformations. For example with $\lambda = 0.5$ equation (9.2) turns into a square root transformation. The effects of various values for λ are given in figure 9.2. It clearly shows that for $\lambda = 1$ the Box-Cox transformation retains the original values, minus 1. For a parameters such as travel time it is expected that λ will be less than 1, as an extra unit of travel time with already large travel times is not as much of influence as that same extra travel time for a small remaining travel time.

The log transformation is a special case of the Box-Cox transformation when $\lambda = 0$. If the continuous variable is lognormally distributed, a log transformation will return a normal distributed variable, with the same mean and standard deviation. For all mentioned continuous parameters a density plot was devised. These are given in figure 9.3.

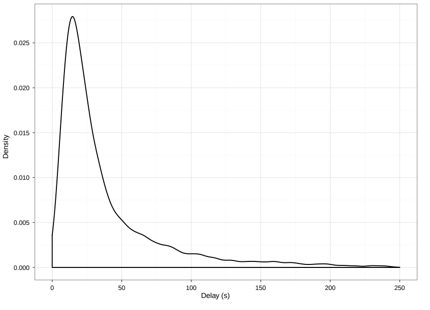
All three parameters are non-normally distributed, where especially the delay has a very high positive skewness and kurtosis, indicating a distribution with a high peak, and skewed towards the lower values. All three are thus appropriate candidates for



(a) *Travel time*



(b) *Freeflow travel time*



(c) *Delay*

Figure 9.3: *Density plots of continuous variables*

a transformation, of which the delay parameter is most closely represented with a log transformation. All three variables are also scaled by multiplying them with 0.01, which means that the estimated values are given for 100 seconds as value of 1. As the log transformation is only valid for values > 0 a value of 1 was used if the variable to be transformed was 0 (in case the log transform was applied).

Approach

The approach in developing a complete exit model start with the investigation of the most appropriate set of attributes of routes, which can be used in the base utility functions. In this step, several combinations of the available attributes are used, together with transformations of the above mentioned non-linear attributes. This step is necessary to investigate the impacts of additional attributes and test for violations of the multinomial logit model, especially that of independent and identically distributed (i.i.d.) unobserved error terms. Firs, the base model is tested for overlap in the routes. This overlap violates the i.i.d. assumption as for the part of routes which uses the same links, the error terms are correlated. Based on the available literature, two corrections for this overlap have been tested, of which the results are presented. The base model is then estimated with additional attributes for accidents, traffic information and guidance on the (relative) importance of these basic route attributes. These attributes show the impact of traffic information and non-recurrence on the route choice behaviour of the participants, and also provides an indication of the (relative) importance of these attributes in relation to those of the base exit choice model. The resulting utility functions together with the assumptions of the multinomial logit model are used to predict the choices for a part of the dataset which is not used of estimation. This is a measure of the internal validity of the estimated choice models. The last step in the approach is to use different modelling assumptions, which relax the restrictive assumptions of the multinomial logit model. These models also test for the assumptions in the exit choice model, such as the independence of previous choices in the same trip.

9.3.3 Exploration of a base model

Basic utility functions

The results for the exploration of an "optimal" basic utility function are given in table 9.8, based on the complete dataset. These models serve as a starting point and a reference for other models. It provides an overview of the attributes which are of importance to the participants in choosing one of the exits at a junction. If no value is provided for an estimated coefficient, the related attribute was not used in the estimation. The two basic combinations that are tested are either based on travel time and delay (TTsimple) or freeflow travel time and delay (FFTTsimple). Since the values of delay are calculated using the difference between travel time and freeflow travel time, it does not contain extra information in the TTsimple model. Accordingly, the final log likelihood and pseudo ρ^2 values are equal. The difference between the models is in the sign of β_D, which is i positive in the TTsimple model. This is inconsistent with normal behaviour, whereas a higher delay would mean that an alternative would be less attractive to a traveller. Next both models also show that freeflow travel time is more important in choosing an alternative than delay. However, the fact that the travel time consists of the freeflow travel time and delay, explains the difference between the two models. The FFTTsimple model would be the more preferred of these two models, because of the sign β_D which allows for an easier interpretation. The ASC's show that for both models, the left exit is not significantly different from the right exit, but the straight exit always has a higher preference. Both the number of junctions and the number of left turns have a negative sign for the estimated coefficient, which is consistent with expectations. A more complex route (more junctions and more left turns) has a smaller probability of being chosen.

The next two models (TTlogD and FFTTlogD) use a log transformation for the delay parameter. In both cases the model shows an improvement in model fit with a lower final log likelihood and higher pseudo ρ^2 compared to the models without transformations. In both cases, the sign of β_D is negative, as could be expected. For the model using travel time, the change in ASC's is small, still rendering the left exit not significantly less preferred than the right exit. The FFTTlogD model however indicates that the left exit does significantly differ from the right exit in preference and has a lower utility and thus probability of being chosen. Both models also show a difference in estimated values for the parameters they both use.

Table 9.8: *Exploration of basis route choice models*

Variable	**TTsimple**	**FFTTsimple**	**TTlogD**	**FFTTlogD**	**logTTlogD**	**BCTT**	**BCFFTT**	**BCTTlogD**	**BCFFTTlogD**
ASC1 (right)	0	0	0	0	0	0	0	0	0
ASC2 (straight)	0.673	0.673	0.771	0.588	0.846	0.737	0.724	0.803	0.619
ASC3 (left)	−0.106[a]	−0.106[a]	−0.073[a]	−0.285	0.010[a]	−0.053[a]	−0.071[a]	−0.036[a]	−0.258
β_{TT}	−0.844	-	−0.437	-	-	-	-	-	-
$\beta_{\log TT}$	-	-	-	-	−3.60	-	-	-	-
$\beta_{TT(\lambda)}$	-	-	-	-	-	−2.86	-	−2.58	-
β_{FFTT}	-	−0.844	-	−0.766	-	-	-	-	-
$\beta_{FFTT(\lambda)}$	-	-	-	-	-	-	−1.99	-	−1.52
β_{D}	0.368	−0.476	-	-	-	0.259	−0.446	-	-
$\beta_{\log D}$	-	-	−0.500	−0.915	−0.370	-	-	−0.325	−0.878
β_{JC}	−0.211	−0.211	−0.377	−0.232	−0.252	−0.125	−0.154	−0.253	−0.191
β_{LT}	−0.131	−0.131	−0.143	−0.100	−0.137	−0.120	−0.123	−0.126	−0.094
λ	-	-	-	-	-	0.375	0.512	0.235	0.620
Number of observations	2517	2517	2517	2517	2517	2517	2517	2517	2517
Null log likelihood	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932
Final log likelihood	−1000.206	−1000.206	−991.241	−949.714	−958.032	−951.617	−986.845	−953.364	−943.821
Likelihood ratio test	2361.451	2361.451	2379.382	2462.436	2445.801	2458.630	2388.147	2455.137	2474.222
Pseudo ρ^2	0.541	0.541	0.545	0.565	0.561	0.564	0.548	0.563	0.567
Adjusted pseudo ρ^2	0.539	0.539	0.543	0.562	0.558	0.560	0.544	0.560	0.564
Number of parameters	6	6	6	6	6	7	7	7	7

[a] Non-significant estimate at a 95% level.

β_D is lower in the FFTTlogD model, indicating a higher influence of delay on the choice for an alternative, with less importance for the amount of junctions or left turns. Of these two models, the FFTTlogD has the best fit, and compared to the two simple models, also has the most sound estimation.

For travel time also a log transformation has been applied, basically restricting λ to 0 in a Box-Cox transformation, which makes it possible to compare it with an unrestricted Box-Cox transformation. This logTTlogD model shows a better fit compared to the TTlogD model, with a higher pseudo ρ^2. The estimates for the other parameters still show the same sign, although the ASC for the left exit is just positive, but still insignificantly different from zero.

For both travel time and freeflow travel time a Box-Cox transformation was applied, where the value of λ was part of the estimation, introducing a new parameter. Again for the model using travel time and no transformation for delay, it shows an inconsistent sign for the estimated coefficient β_D, whereas the BCFFTT model does show the appropriate sign. This indicates the same problem as in the TTsimple model, where delay is double counted in the model. For the ASC's, the values are consistent with the previous models, with the straight exit still having the largest intricate preference. Both estimates for number of junctions and number of left turns are also consistent with previous estimates, although the are showing less difference in value. The value for λ is in both models significantly different from both 1 and 0, indicating that it is neither linear nor log-normally distributed, where the value of λ for the BCFFTT model is slightly higher. This indicates that an extra unit of travel time is slightly less of impact compared to an extra unit of freeflow travel time. Interestingly, of both BCTT and BCFFTT models, the first has the best fit, even though the inconsistent estimate of β_D.

The last two estimated models combined the Box-Cox transformation for travel time or freeflow travel time with the log transformation of delay. Of these two models, the BCFFTTlogD model has the best fit of all estimated models (highest pseudo ρ^2 and lowest final log likelihood). Both models show estimated coefficients which are consistent with respect to the signs and values with previous models, whereas the BCFFTTlogD model also shows a significant value of the ASC of the left exit (less preferred compared to the right exit). Both models have estimates for λ that are significantly different from 0 and 1. For the amount of left turns the BCFFTTlogD model has the lowest value, but it is still significantly different from zero, and rather close to the value estimated in the FFTTlogD model. The same applies to the other estimated coefficients. This indicates that the Box-CoX transformation adds extra explanatory value to the freeflow travel time, improving the model fit. This is also shown using the Likelihood ratio test of the FFTTlogD and BCFFTTlogD model, which shows a significant difference.

Based on these explorations a base utility function can be set, including the following attributes:

- (Box-Cox transformation of) freeflow travel time (s/100);
- (log transformation of) delay (s/100);
- amount of junctions, and;
- amount of left turns.

Table 9.9: *Exploration of additional parameters using the base models*

Variable	Complete dataset *Base*	Complete dataset *Additional*	Strategic dataset *Base*	Strategic dataset *Additional*
ASC1 (right)	0	0	0	0
ASC2 (straight)	0.619	0.587	−0.167[a]	−0.211[a]
ASC3 (left)	−0.258	−0.192[a]	−0.306[a]	−0.257[a]
$\beta_{FFTT(\lambda)}$	−1.52	−1.08	−3.36	−3.51
$\beta_{\log D}$	−0.878	−0.780	−0.754	−0.617
β_{JC}	−0.191	−0.273	−0.128	−0.241
β_{LT}	−0.094	−0.075[a]	0.042[a]	0.079[a]
β_S	-	−0.011[a]	-	−0.015[a]
β_C	-	0.601	-	0.360[a]
λ	0.620	0.597	0.221[a]	0.023[a]
Number of observations	2517	2517	905	905
Null log likelihood	−2180.932	−2180.932	−800.432	−800.432
Final log likelihood	−943.821	−933.842	−480.771	−477.824
Likelihood ratio test	2474.222	2494.181	639.321	645.216
Pseudo ρ^2	0.567	0.572	0.399	0.403
Adjusted pseudo ρ^2	0.564	0.568	0.391	0.392
Number of parameters	7	9	7	9

[a] Non-significant estimate at a 95% level.

Expanded basic utility functions

Using this base function, two additional parameters are investigated, using both the complete and strategic dataset. The additional parameters are the average speed and a dummy for being closer to the destination as the crow flies. The results are given in table 9.9.

First, when comparing the functions for the complete dataset, the Likelihood ratio test shows a significant improvement, although the model fit is not much better. The addition of the extra two parameters average speed and "closer" renders the ASC for the left exit and the coefficient for the number of left turns β_{LT} insignificant. The estimated coefficient for speed is also insignificant, which is likely because the average speed of an alternative is expected to be correlated to the freeflow travel time. The coefficient for being closer to the destination β_C however is significant and has a positive sign, which is consistent with expectations. As the alternative brings the traveller closer to the destination, that alternative is more likely to be chosen than an alternative which brings the driver further away, which is a likely response of most drivers. The addition of the two new parameters also causes a reduction of the value for the coefficients of delay and freeflow travel time. The value of λ is more or less equal. The estimated coefficient for amount of junctions is increased, making it more important. In combination with the other estimates, it indicates that because being closer is important, the complexity of a route is also becoming more important, whereas the values for delay or freeflow travel time are more or less equal.

Both utility functions have also been applied to the dataset with the strategic junctions. This results in a lower pseudo ρ^2 compared to the models fitted with the complete dataset and the likelihood ratio test also shows lower values, although this indicates that all estimated coefficients together are still significantly different from zero. When looking at the ASC's the results indicate that neither the straight or the left exit has a significantly different preference from the right exit. This differs from the complete dataset which indicates a preference for the straight exit. The locations of the strategic junctions are thus influencing the choice of participants, removing any preference for a certain type of exit. In both models, four other coefficients show an estimate with significant value, with the λ not significantly different from zero, but from one. The estimates for average speed and "closer" are not significantly different from zero, which makes the additions not provide much better model fit, which is also given by the likelihood ratio test using both models ($p>0.05$). Because the λ value is not significantly different from zero, the Box-Cox transformation applied to the freeflow travel time more or less reduces to the log transformation. Both models show consistent signs for the significant coefficients. Apart from the value for freeflow travel time, the values for delay and amount of junctions are in the same range as in the models estimated with the complete dataset. With the coefficients for delay and amount of junctions more or less equal with the complete dataset estimates, it indicates that at the strategic junctions, the freeflow travel time of the alternatives is more important, regardless of the specific exit.

The base models also have been estimated using alternative specific coefficients, such as β_{D_i}. Although the estimated coefficients varied slightly, using a t-test showed that neither of these alternative specific coefficients are significantly different from each other. Together with the increased number of parameters, the use of alternative specific coefficients is not used.

The results show that the proposed exit choice model is possible to estimate and provides estimates for a basic set of attributes which are reasonable and have a relative difference and sign as might be expected. The assumed MNL model used for the estimation already provides results with a high model fit. The two resulting utility functions are equation (9.3) for the complete dataset and equation (9.4) for the strategic dataset.

$$V_i = ASC_i + \beta_{FFTT} FFTT_i(\lambda) + \beta_{logD} \log D_i + \beta_{JC} JC_i + \beta_C C_i \tag{9.3}$$

$$V_i = ASC_i + \beta_{FFTT} FFTT_i(\lambda) + \beta_{logD} \log D_i + \beta_{JC} JC_i \tag{9.4}$$

In these models, freeflow travel time is the most important attribute, which is transformed using a box-cox transformation. With longer travel times, an additional unit of travel time is of less influence on the choice, compared to shorter travel times. This is reasonable behaviour, as for a short trip, a small difference in freeflow travel time is considered to be more noted, then for a long trip. Travel time (both in terms of delay and freeflow) is always decreasing the utility of an alternative, which means the participants prefer shorter and quicker routes with less delay. Less complex routes, with less choices (junctions) means the route is more likely to be chosen by the participants. The direction of the exit is also an important attribute, which indicates that participants prefer routes that brings them towards the destination. Especially in case of accidents on this route, it would be interesting to see if this preference holds. However, it is necessary to investigate the assumptions underlying both the MNL model, and the proposed exit choice model, which is presented in the next section.

9.3.4 Additional effects

Correcting for overlap

Using the exits as the alternatives in the choice set did mean one of the basic assumptions of the MNL model is violated. As the alternatives still have overlap (in terms of links and thus length and travel time), the error or unknown terms are not independent (see also chapter 4. This overlap leads to correlations, which can significantly influence choice probabilities, as discussed in Bovy et al. (2008). Of the available definitions for the correction factors, just two (Path Size and Path Size Correction, see section 4.3.3) have been used, as these have been shown to perform the best (Ramming, 2001; Prato & Bekhor, 2007). The two factors are repeated here for convenience.

$$PS_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \frac{1}{\sum\limits_{j \in \mathcal{C}_n} \delta_a j} \tag{9.5}$$

$$PSC_{in} = \sum_{a \in \Gamma_i} \frac{L_a}{L_i} \ln \left(\frac{1}{\sum\limits_{j \in \mathcal{C}_n} \delta_a j} \right) \tag{9.6}$$

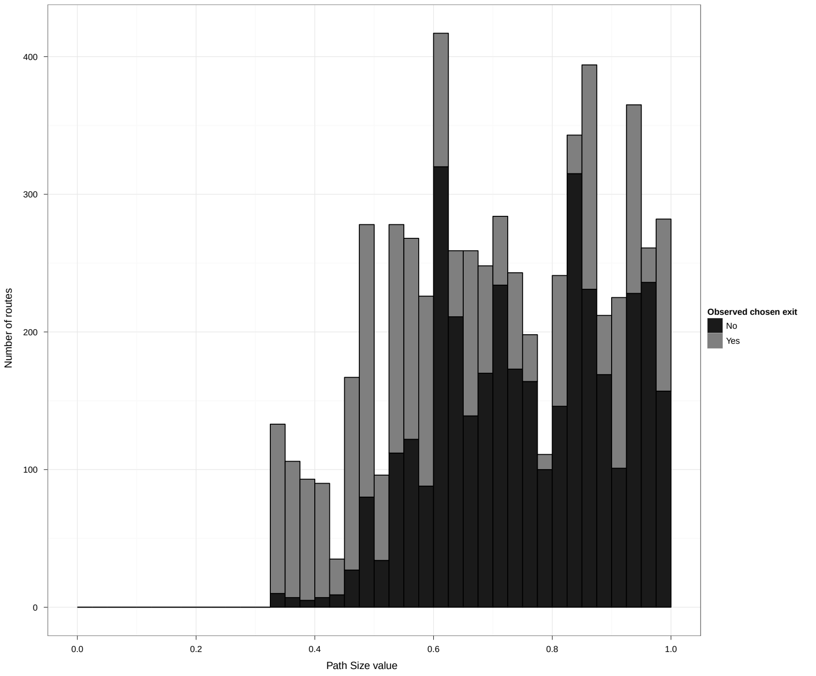
The L_a and L_i are defined as length in metres in this research (for example as in Frejinger (2008)), but could also be used with values for freeflow travel times (for example as in Ramming (2001)). The total linear-in-parameters utility equations are as given in equations (9.7) and (9.8) in which the subscript for the individual n is dropped. V_i refers to the utility function given in equation (9.3) and equation (9.4).

$$V_i' = V_i + \beta_{PS} \ln PS_i \tag{9.7}$$

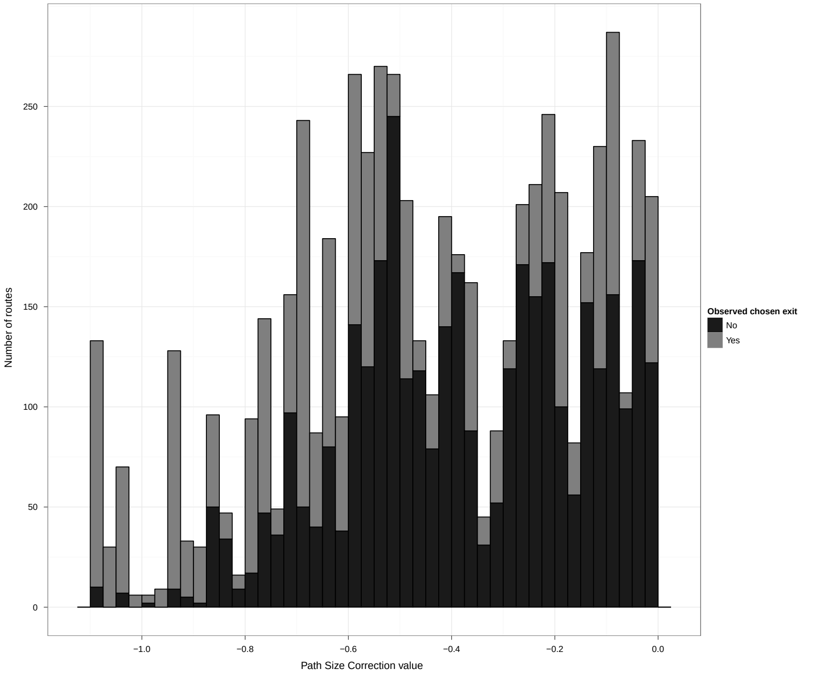
$$V_i' = V_i + \beta_{PSC} PSC_i \tag{9.8}$$

In this specific application there are either two or three routes available, depending on the type of junction where a choice is made. Because the choice set is created using all the available exits at a junction, the overlap or commonality of the two or three routes never occurs at the start of the route. Next, the routes always lead to a single destination, which is only reachable using a single link. This causes all routes to always have at least one link in common at the end of the route. The values for both the Path Size and the Path Size Correction factor are displayed in figure 9.4, where a distinction is made for the values for routes that are chosen and the other routes. The histograms are providing the stacked values.

The Path Size (PS) values range from 0.333 to 0.994, with a mean of 0.711. A PS value of 1 would indicate an unique route, with a PS value of 0 a route with complete overlap. The resulting histogram is quite different from those presented in (Frejinger, 2008; Schüssler, 2010), where both have values for the complete range of PS values. This is due to the fact that the choice set in this approach is fairly small



(a) *Path Size*



(b) *Path Size Correction*

Figure 9.4: *Histograms of overlap factors for the complete choiceset*

and using the different exits at a junction. For lower PS values, the histogram indicates that routes with a higher overlap are chosen more often, compared to higher values.

The Path Size Correction (PSC) values range from -1.099 to -0.008 with a mean of -0.438. The values of PSC are negative because it already contains the ln transformation. A PSC value close to zero indicates a route that is unique, whereas the PSC values approaches $-\infty$ the overlap of the route increases. Again, the histogram indicates that routes with more overlap are more often chosen.

For both the Path Size and Path Size Correction factors an extra parameter can be estimated, although theory suggest these parameters should be 1, as these are just correction factors. However, as previously suggest by (Hoogendoorn-Lanser, 2005; Frejinger, 2008) the estimated parameters (β_{PS} or β_{PSC}) also have a behavioural interpretation. They capture the attractiveness of the overlap of the routes, which would allow for the possibility to switch routes. This explanation however is not valid for this choice set, as the overlap always occurs at the end of the routes, which does not relate to the possibility to switch. In this case it is more likely to concern routes that are intrinsically faster compared to all others, which might explain any behavioural implications of the correction factors.

The results for the estimated models for both datasets are given in tables 9.10 and 9.11. For both datasets, the addition of the correction factor (either as Path Size or Path Size Correction) have a similar effect on the other parameters. In both cases the estimated values for $\beta_{FFTT(\lambda)}$ and λ are significantly different compared to the model without the correction factors. The value for freeflow travel time is reduced, indicating a much smaller impact on the (dis)utility of the route. In the case of the complete dataset λ is increased to a value close to one, and it is no longer significantly different from one. The Box-Cox transformation for freeflow travel time is therefore no longer necessary when applying the correction factors. In case of the strategic dataset, λ is both insignificant from one and from zero. This also renders the Box-Cox transformation of freeflow travel time not applicable to this model, whereby the reasonable approach would be to use the unity values. The other variables in both datasets are in the same range and have consistent signs. This indicates a big explanatory value of freeflow travel time and the added correction factors.

For both datasets and both correction factors, the estimated correction coefficients have a negative sign. An increase in overlap would accordingly decrease the disutility of an alternative (for Path Size a lower value indicates more overlap, which becomes more negative after the log transformation, the same applies to the Path Size Correction factor although the log is already used in the calculation of the factor). As mentioned before, this would indicate that travellers have a preference for route with more overlap, which is inconsistent with the (theoretical) derivation of the Path Size Correction factor.

The other possible explanation that the increase in overlap would allow for travellers to prefer that alternative is caused by the possibility to switch to other alternatives later, is difficult to apply in this case. First of all, the overlap in the routes is always located at the end of each alternative, caused by the use of the possible

Table 9.10: *Base models with the addition of a correction factor for the complete dataset*

Variable	Base	PS	PSC
ASC1 (right)	0	0	0
ASC2 (straight)	0.539	0.510	0.512
ASC3 (left)	−0.285	−0.285	−0.280
$\beta_{FFTT(\lambda)}$	−1.18	−0.489	−0.467
$\beta_{\log D}$	−0.859	−0.763	−0.770
β_{JC}	−0.223	−0.189	−0.198
β_{C}	0.611	0.674	0.662
β_{PS}	-	−1.70	-
β_{PSC}	-	-	−1.89
λ	0.637	0.971	0.991
Number of observations	2517	2517	2571
Null log likelihood	−2180.932	−2180.932	−2180.932
Final log likelihood	−937.439	−922.219	−922.934
Likelihood ratio test	2486.987	2517.427	2515.996
Pseudo ρ^2	0.570	0.577	0.577
Adjusted pseudo ρ^2	0.567	0.573	0.573
Number of parameters	8	9	9

[a] Non-significant estimate at a 95% level.

exits at a junction. This eliminates the possibility of having overlap at the start of an alternative. The algorithm that calculates the routes also reduces the possibility that there is a diversion between the alternatives further on the route to zero, because it used the same information for its' wayfinding (the current travel times). Second, the choiceset at each junction consists of two or three alternatives. This amount of alternatives is very low compared to other studies, and the size of the choice set is assumed to be fairly large in the derivation of the correction factors. The choice set does have a great influence on the correction factors, and the choice modelling in general (see for instance (Schüssler, 2010; Prato & Bekhor, 2007; Bliemer & Bovy, 2008). Next, as the alternatives are labelled, the systematic utility is not necessarily equal (as for example, the alternative specific constant differs). Third and last, the usual application of the correction factor is in a single route choice between origin and destination. In this case, the route choice is evaluated at all locations where a choice was possible. If the choice is then related to just the exit at a junction, overlap is never possible for these exits. The possible explanation of having a route that is intrinsically faster than all other alternatives still holds. This also explains the overlap with the freeflow travel time attribute.

Although the likelihood ratio test for both datasets and both correction factors indicate a significant model improvement ($p<0.05$) and have higher pseudo ρ^2, for the investigation of the effects of accidents, traffic information and route guidance, the correction factors are not applied. Also just using the value of the correction factor, without estimating a coefficient is not used, as these models provided a worse model fit compared to the base models. This choice is based on the difficult explanation for the coefficients of the correction factors, and the possible overlap with the freeflow travel time. It is preferred to use the freeflow travel time over the correction factors.

Table 9.11: *Base models with the addition of a correction factor for the strategic dataset*

Variable	**Base**	**PS**	**PSC**
ASC1 (right)	0	0	0
ASC2 (straight)	-0.110^{a}	-0.115^{a}	-0.151^{a}
ASC3 (left)	-0.244^{a}	-0.270^{a}	-0.272^{a}
$\beta_{FFTT(\lambda)}$	-3.43	-1.70^{a}	-1.80^{a}
$\beta_{\log D}$	-0.745	-0.672	-0.686
β_{JC}	-0.122	-0.102	-0.110
β_{PS}	-	-1.43	-
β_{PSC}	-	-	-1.42
λ	0.197	0.513^{a}	0.489^{a}
Number of observations	905	905	905
Null log likelihood	-800.432	-800.432	-800.432
Final log likelihood	-481.125	-475.975	-477.264
Likelihood ratio test	638.613	648.914	646.336
Pseudo ρ^2	0.399	0.405	0.404
Adjusted pseudo ρ^2	0.391	0.397	0.395
Number of parameters	6	7	7

[a] Non-significant estimate at a 95% level.

Incidents, traffic information and guidance

As a next step, the effects of incidents, traffic information and route guidance on the choice for an exit are studied. These parameters are dummies, indicating if either of these parameters is available or exists for a specific alternative. If traffic message is valued as one, it indicates that for this exit and the route that belongs to that exit, a traffic jam is reported to the traveller for one of the links of that route. The same applies to the dummy for the incident. Advice is given as one, when the alternative or exit is being advised to the traveller based on the fastest route from that location.

For these three variables the amount of choices with respect to the two datasets and choices are given (see table 9.12). For example, in the strategic dataset, 905 choices were made for an exit with a route where no traffic message was provided for (and thus no traffic jam existed on that route). For traffic messages, a possible problem arises, as for the strategic locations, there is no exit chosen with a route which had a traffic message about a traffic jam. This makes it difficult to estimate a coefficient for this variable, as no data is available for this combination. When looking at the total amount of traffic messages for all exits, regardless if these are chosen, it sums up to 85, compared to 6027 alternatives without a traffic message. This amount is fairly low, which is also a result of the applied route generation algorithm.

For incidents, more choices are available in both the complete and strategic datasets. In total 150 alternatives were generated which contained a link which had an incident at that time. Interestingly, the route generation algorithm created routes that were fastest for that exit, but still passed the location of an incident, and in a number of cases (44) this alternative was also chosen by the traveller. The parameter of advice contains the most data for positive values (advice was available). It also shows that in case an advice was available (just in two trips per participant), the

Table 9.12: *Contingency tables for parameters of the alternatives*

(a) *Traffic messages*

		Strategic	
Traffic message	**Chosen**	*No*	*Yes*
No	*No*	2228	1285
	Yes	1609	905
Yes	*No*	35	47
	Yes	3	0

(b) *Incidents*

		Strategic	
Incident	**Chosen**	*No*	*Yes*
No	*No*	2250	1239
	Yes	1595	878
Yes	*No*	13	93
	Yes	17	27

(c) *Advised*

		Strategic	
Advised	**Chosen**	*No*	*Yes*
No	*No*	2238	1307
	Yes	1218	696
Yes	*No*	25	25
	Yes	394	209

advised exit is quite often selected (92,3% of all cases). For the strategic dataset, the choice for the advised exit reduces to 89,3%.

The results of the estimation of the newly added coefficients are presented in table 9.13, where the base models are included for reference. In case of the complete data set, the results provide a significantly better fit with the model including the added effects, based on the likelihood ratio test ($p<0.05$). Also the pseudo ρ^2 is higher, compared to the base model. The addition of the extra parameters for traffic messages, incidents and advice thus provide extra information about the choice for an exit. The same applies to the strategic dataset, where the pseudo ρ^2 also is higher compared to the base model. The likelihood ratio test also shows a significant improvement ($p<0.05$).

The complete dataset allowed to estimate a coefficient for all three parameters. All of the new coefficients have the appropriate sign, with the coefficient for the alternative that is advised having a positive value, as this alternative is based on the fastest route at that choice. The values for incidents and traffic messages are both negative, with the value for incidents not significantly different from zero ($p>0.05$), but this parameter will still be used in the model, as it shows the effect of incidents. The value of the coefficient for traffic messages is large compared to the others, but this is most likely due to the small amount of available traffic messages and choices for alternatives with a traffic message. The value of advice is as expected also rather large, which can be explained by the high number of choices for exits which were advised. The other coefficients are in the same range as in the base model, apart from delay. Although still significantly different from zero, this value is smaller. This indicates that the addition of the new parameters add a lot of explaining value to the model, part of which was explained by the delay. Regarding the ASC's, the ASC for the left exit is reduced as well, and became insignificantly different from zero.

Table 9.13: *Base models with the addition parameters for incidents, traffic messages and advise*

Variable	Complete dataset *Base*	Complete dataset *Effects*	Strategic dataset *Base*	Strategic dataset *Effects*
ASC1 (right)	0	0	0	0
ASC2 (straight)	0.539	0.577	−0.110[a]	−0.049[a]
ASC3 (left)	−0.285	−0.194[a]	−0.244[a]	−0.040[a]
$\beta_{FFTT(\lambda)}$	−1.18	−1.05	−3.43	−3.25
$\beta_{\log D}$	−0.859	−0.574	−0.745	−0.491
β_{JC}	−0.223	−0.237	−0.122	−0.148
β_C	0.611	0.597		
β_I		−0.618[a]		−0.778
β_{TM}		−3.88		
β_A		1.39		1.28
λ	0.637	0.633	0.197[a]	0.138[a]
Number of observations	2517	2517	905	905
Null log likelihood	−2180.932	−2180.932	−800.432	−800.432
Final log likelihood	−937.439	−878.793	−481.125	−458.174
Likelihood ratio test	2486.987	2604.277	638.613	684.516
Pseudo ρ^2	0.570	0.597	0.399	0.428
Adjusted pseudo ρ^2	0.567	0.592	0.391	0.418
Number of parameters	7	10	6	8

[a] Non-significant estimate at a 95% level.

The model applied to the strategic dataset with the additional effects does not include the parameter for traffic messages. As mentioned, no data is available for this parameter combined with a choice for the alternative containing a traffic message. This renders the model unidentifiable. Therefore, just the incidents and advice are used for the strategic dataset. Both of these coefficients have a consistent sign, and the values are also in the same range as for the complete dataset. This is different from the ASC's and the coefficients for freeflow travel time which had a different value between both datasets. The specific location is thus less of influence on the size of the added effects. The coefficient for incidents is significant in this model and more important than the delay. Freeflow travel time still is the most important parameter in this model with the highest value.

The resulting utility functions for both datasets are given in equations (9.9) and (9.10), which will be used in further analysis.

$$\begin{aligned} V_i = & ASC_i + \beta_{FFTT} FFTT_i(\lambda) + \beta_{logD} \log D_i + \beta_{JC} JC_i + \beta_C C_i \\ & + \beta_I I_i + \beta_A A_i + \beta_{TM} TM_i \end{aligned} \tag{9.9}$$

$$\begin{aligned} V_i = & ASC_i + \beta_{FFTT} FFTT_i(\lambda) + \beta_{logD} \log D_i + \beta_{JC} JC_i \\ & + \beta_I I_i + \beta_A A_i \end{aligned} \tag{9.10}$$

Table 9.14: *Probability of selecting an alternative aggregated for observations not used for estimation*

Exit	Predicted	Observed
Right	0.183	0.200
Straight	0.631	0.616
Left	0.185	0.183

9.3.5 Prediction and internal validity

One of the possible applications of a discrete choice model is to predict choice probabilities. To this end, the utility function with previously estimated coefficients is evaluated for a specific set of values for the variables of this utility function. This results in a value of the utility for a specific alternative, which is used in the MNL model to predict the probability that the alternative is chosen. This allows one to predict market shares, or the in this case, the probability a driver will select a left, right or straight exit at a junction.

It also allows to test for the internal validity of a utility function and discrete choice model. To this end, a part of the dataset is used for estimation of the coefficients in the utility function, using a number of assumptions for the choice model, such as for the MNL model. The resulting utility function can then be applied to predict the probability of the alternatives for the remainder of the dataset (the part not used for estimation). These probabilities can be compared with the observed probabilities. This comparison is a measure of the internal validity of the model used.

This approach has been applied in two ways. The first excluded every 8th choice in the complete dataset, which resulted in a total of 2212 observations which were used for estimation of the utility function in equation (9.9). The remaining 305 choices were used for a prediction, which was compared with the observed choice for an alternative. The second approach estimated the same utility function for each scenario, applying it to the other two scenarios for prediction.

Prediction of ⅛ of the observations

The estimated coefficient for the first approach, using ⅞ of the observations, showed that reducing the amount of observations does not have a large influence on the estimated parameters or the model fit. The estimated values for pseudo ρ^2 are even equal to the model with all observations. The only real difference is in the significance of having an incident on the route of the alternative, which is significantly different from zero in this estimation.

Using the estimated parameters, the choice for the remaining observations are reproduced. The probability of choosing a certain alternative almost perfectly matches the percentage of observations of these choices. The resulting probabilities are given in table 9.14. As the model fit is very good, since the high pseudo ρ^2, it is not surprising that the prediction of the unused observations is good.

Table 9.15: *Base model estimated on each scenario*

Variable	**Scenario 2**	**Scenario 3**	**Scenario 4**
ASC1 (right)	0	0	0
ASC2 (straight)	0.913	−0.066[a]	0.818
ASC3 (left)	−0.411[a]	−0.914	−0.083[a]
$\beta_{FFTT(\lambda)}$	−3.56	−0.891	−1.32
$\beta_{\log D}$	−0.234[a]	−0.322	−0.604
β_{JC}	0.348	−0.121[a]	−0.305
β_C	0.824	1.47	−0.336[a]
β_I	−1.48	−1.59	2.01
β_{TM}	−5.16[a]	−25.6[a]	−3.61
β_A	2.11	0.667	1.79
λ	0.450	0.209	0.924
Number of observations	732	680	1083
Null log likelihood	−673.061	−574.734	−933.138
Final log likelihood	−180.075	−258.562	−365.571
Likelihood ratio test	985.971	632.343	1135.133
Pseudo ρ^2	0.732	0.550	0.608
Adjusted pseudo ρ^2	0.718	0.533	0.598
Number of parameters	10	10	10

[a] Non-significant estimate at a 95% level.

Prediction of the scenarios

Table 9.15 shows the resulting estimated coefficients for the different scenarios. The estimated coefficients are quite different between the three scenarios. The results suggest that the combination of origin, destination and viable routes have an important role in the estimates of the route choice model, as a different scenario provides different results for the estimated coefficients. For example, scenario 3 shows no significant difference between the alternative specific constants for straight ahead and right exits, which is different from both other scenarios where the straight ahead exit is preferred over the right exit. The three scenarios also show different estimates for the freeflow travel time parameter and the parameter for the Box-Cox transformation. In case of scenario 4, there is hardly any non-linearity in the freeflow travel time, whereas scenario 3 is the opposite. However, the transformed value of freeflow travel time is the smallest in scenario 3, which means in this scenario an effect of an additional unit of freeflow travel time is larger for low values of freeflow travel time compared to the other scenarios. However, as the value of freeflow travel time is the lowest of all scenarios, it is likely to have less effect on route choice in general.

There are also major differences in the additional parameters. In scenario 4, incidents have a positive sign, which is hardly reasonable. This might be caused by a limited amount of choices for an exit with incidents on the route. The same applies to the number of junctions, which has a positive value in scenario 2. Also interesting to note is the difference in the estimated coefficient for advice, which is lowest in scenario 3. This is most likely due to the effect of the north-south orientation of this scenario which caused the given advice to be trusted less by the participants. Scenario 3 also has the smallest value of pseudo ρ^2, which means it has got the worst model fit of the three models, albeit still rather good.

Table 9.16: *Probability of selecting an alternative aggregated for observations not used for estimation*

	Scenario 2		Scenario 3		Scenario 4	
Exit	*Predicted*	*Observed*	*Predicted*	*Observed*	*Predicted*	*Observed*
Right	0.158	0.176	0.303	0.213	0.205	0.238
Straight	0.667	0.630	0.561	0.636	0.653	0.614
Left	0.175	0.194	0.136	0.151	0.141	0.149

The three separately estimated models for the three scenarios have been used to simulated the choices for the remaining scenarios. For the estimated model using scenario 2 it means that scenario 3 and 4 have been used for simulation, etc. The resulting predicted probabilities and observed probabilities are presented in table 9.16.

The predicted probabilities from scenario 2 and 4 applied to the unused scenarios are closest to the observed probabilities. Especially scenario 4 is able to reproduce most of the choices in the other scenarios, with an overestimation of straight ahead and underestimation of right exits. Left exits are quite close with just 0.008 difference in probability. The results of scenario 2 show some bigger differences between prediction and observation, but still is quite able to reproduce the choices. Scenario 3 however, as expected, has more trouble reproducing the choices from the other two scenarios. The straight ahead exits are underestimated with a difference of 0.075, which is quite a lot. The right exits however are strongly overestimated by 0.09. These results indicate that scenario 3 is different regarding the participant choices compared to the other scenarios. This was also indicated by the different scale parameter for scenario 3, compared to scenario 2 and 4. The results of scenario 3 are thus not easily transferable to a different combination of origin and destination. Scenarios 2 and 4 however, are capable of reproducing choices in other combinations of origin and destination. The internal validity of these two scenarios is thus quite high, whereas it also could indicate the results of these scenarios are applicable in other cases as well.

9.3.6 Relaxing the MNL assumptions

Using both base models from equations (9.9) and (9.10), various other explorations have been made which try to relax the assumptions made for the MNL model, regarding the unobserved utility or error terms or the assumptions of the exit choice model (as a sequence of choices is unlikely to be independent). In the previous estimated models, these assumptions, especially that of independent and identical distribution of the unobserved or error terms, are most likely violated. The following models are each used to investigate the degree in which these assumptions are violated. The first model investigates the assumption that the unobserved terms in the utility function have the same variance for different segments in the dataset (identical distribution), by estimating different scale parameters for each segment.

The assumption that the unobserved utility is independent between alternatives, (consecutive) choices and decision makers is also investigated. This assumption is violated as each participant made a sequence of trips and within a trip a sequence of choices, for which it cannot be assumed they are independent. Two different approaches were used to relax this assumption: error component logit with normalization of one of the error components and error component logit with independent errors over alternatives and participants. Fourth, the heteroscedastic logit model has been applied, which resembles the error component logit with normalization, but ignores the sequence of observations. As such, it assumes different variances for the alternatives. Fifth and last, the assumption that each estimated coefficient in the utility function has the same value for each observation is relaxed, by assuming a distribution for these coefficients. For this distribution, the parameters are estimated, in case of a normal distribution the mean and standard deviation. This is called random parameter logit.

Scale parameters

Using these segmentations allows to estimate the scale parameter μ, with one of them normalized to 1. Recall that the variance of the Extreme Value Type 1 distribution is given by $\mu^2(\pi^2/6)$. The scale is thus a measure of variance. As given in section 4.2.1, it is possible to estimate a different scale parameter for different segments, compared to an arbitrarily set scale parameter of one for one segment, as long as the utility functions are linear in parameters. The difference in scale parameters then indicates a difference in variance of the segments over all the coefficients in the utility function, regardless of the alternative. While this is most often used in combining stated preference with revealed preference data, it is easily applied to this dataset. The scale parameters show how various segments in the dataset are showing the same variation in their choices. As such, it is an indication of the homogeneity of the different elements in the experiment, such as scenarios, accidents and participants.

A number of segments have been defined, using the gender, the individual participant, the trip of the participants, the specific scenario, the index of the incident and the type of information provided, all for the complete and strategic dataset. If the scale parameters that are estimated are significantly different from 1 (the normalized scale), it indicates a significant difference in variance of choice making between these two groups. The estimated coefficients for the complete model are provided in appendix F for reference, whereas the estimated scale parameters are presented here (tables 9.17 and 9.18). The results of the scale parameters for all individual participants have been omitted here, but can also be found in appendix F.

In case of the complete dataset, often the scale for a group is found to be insignificant from one. This basically indicates the variance of the segments are not significantly different. In the case of the scenario's, scenario 2 was taken as the reference group. The scale parameter for scenario 3 (indicated by a) is significantly lower. A lower scale means a larger variance, which is easily verified with the amount of "problems" participants had in this scenario because of the north-south route between the origin and destination.

Table 9.17: *Estimated scale parameters for various segments (complete dataset)*

Scale	Scenario	Incident	Information	Trip	Gender	Age	License
a	0.617	1.84	1.47	2.01	1.15[a]	0.581	0.663
b	1.01[a]	1.31[a]	1.47[a]	1.47			
c		0.798[a]		1.33[a]			
d		1.21[a]		2.26			
e		1.49		1.83			
f		1.25[a]		1.79			
g				2.53			

[a] Non-significant estimate at a 95% level.

In case of incidents, the trips without an incident, regardless of the scenario, were taken as the reference group. Only two specific incidents show a significant scale parameter compared to the trips without incidents: incident 1 and incident 5. With a larger scale parameter, this indicates a smaller variance or in other words less difference in route choice between the participants. When looking at the images of the trips for incident 1 (figure 8.2a) and incident 5 (figure 8.2e), this finding can be confirmed, because almost all participants made the same route choice in these circumstances.

The scale parameter for having textual information is also significantly different from one, where the reference group is the trips without information. A larger scale parameter also means that in this case the variance is reduced between the trips made by the various participants. Interestingly, the scale parameter for the trips made with route guidance is not significantly different from one, which means that the same variance exists for the trips with route guidance and no information. This finding is not confirmed by the other findings, such as the percentage of choices which are the same as the advice.

The results for the specific trip made by a participant (aggregated over all participants, such that trip 1 are all first trips, trip 2 are all second trips, etc.) show a large increase of scale parameters for all but one trip. In all cases the variance thus is reduced compared to the first trip made by the participants. This indicates a very fast learning effect, which however is not consistent for all trips, as there is no trend in the increase of the scale parameter. Since these scales are for groups aggregated over the participants, it is confounded with information, which varied in a consistent manner over the trips, whereas the specific scenario defined the incident. As such, it does indicate that the last two trips have a smaller variance in the various trips.

When looking at the effects of gender, there is no significant difference between men and women (men were used as the reference group). Even though, the positive value is interesting as it indicate that women agree more in their route choice. Age as a segmentation shows a significant difference between both groups, where the participants aged below 30 are used as a reference. Older participants have a smaller scale parameter and thus exhibit more variance in their route choice. The amount of years possessing a divers' license also shows a significant difference for the participants which possess a drivers' license for more than 10 years. This indicates the same effect as age, since one has to be older to be able to possess a drivers' license over 10 years. The results for participants individually also indicate a few

Table 9.18: *Estimated scale parameters for various segments (strategic dataset)*

Scale	Scenario	Incident	Information	Trip	Gender	Age	License
a	0.706	1.19[a]	1.94	1.32[a]	1.12[a]	0.602	0.714
b	1.15[a]	1.40[a]	1.71[a]	0.930[a]			
c		0.762[a]		0.995[a]			
d		1.19[a]		2.02[a]			
e		2.05		2.04			
f		1.08[a]		1.94[a]			
g				1.72[a]			

[a] Non-significant estimate at a 95% level.

users having a significantly different scale compared to the first participant. In all of these cases, the scale parameter is lower, indicating a higher variance. When combining this with the experiment notes, it is verified that these participants had indeed difficulty making choices during a few trips. Not surprisingly, this applies to most that participated in scenario 3.

Table 9.18 shows the same scale parameters for the strategic dataset. Apart from the scale parameters for the trip indicator, there are not much differences between the two datasets. Variance is thus not directly related to the location of the choices for which the scales are calculated, apart from the number of the specific trip. In case of the complete dataset, participants showed less variance in the later trips, compared to the first trip, whereas in the case of the strategic dataset, only the 6^{th} trip shows a significant decrease in variance. However, even though the individual scale parameters show insignificance from one, the later trips, especially the four where traffic information or route guidance was provided, have a larger value, compared to the previous trips. Besides, the two trips (3 and 4) during which no information was provided show a (albeit tiny) decrease in scale compared to the first trip. The direction of change in the scale parameter thus does imply a certain relation to the circumstances under which a trip was made, and that this had an effect on the route choice of the participant in a certain trip.

Error component logit and repeated observations

As in the survey for the preferences for traffic information (see chapter 5), the route choice experiment had multiple responses from one participant, essentially making the data that of a panel. As already described in the segmentation, using the specific index of the trip a participant made, as well as the participant itself gave indications of having unequal variances between participants and between the trips. However, this does not mean the sequence of choices of the same participant has been accounted for. As described in section 4.2.5, the mixed logit approach enables a researcher to account for multiple choices by the same participant. In this case, three definitions of multiple or repeated choices have been used, which all have different assumptions about the correlation between observations. These three are:

1. the participant (UserID);
2. the index of the trip made by each participant (UserTripID), and;
3. the index of the individual trip (FullTripID).

The first indicates 30 different individuals, the second 8 different trips, and the last gives 269 different trips. The last one essentially takes a single route as a set of correlated observations, and is closest to a "normal" route choice model.

Two different approaches were used to apply these three definitions. Both approaches use the standard normal distribution. In the first approach (also given in equation (9.11), the distribution is independent over the three alternatives, where the alternative with the lowest variation is normalized for identification (see Walker (2001); Walker et al. (2007)).

$$\begin{aligned} U_R =& V_R + \sigma_R \xi_R + \varepsilon_R \\ U_S =& V_S + \varepsilon_S \\ U_L =& V_L + \sigma_L \xi_L + \varepsilon_L \end{aligned} \tag{9.11}$$

In the second approach (described in Hess & Rose (2006); Hess et al. (2008), see equation (9.12)), the standard normal distribution is used, but is assumed to be independent over the alternatives and participants, but not across observations of the same respondent. This is different from the previous approach, as it does not create correlation between some of the alternatives which happens in the previous approach (see also Yáñez et al. (2010)).

$$\begin{aligned} U_R =& V_R + \sigma \xi_R + \varepsilon_R \\ U_S =& V_S + \sigma \xi_S + \varepsilon_S \\ U_L =& V_L + \sigma \xi_L + \varepsilon_L \end{aligned} \tag{9.12}$$

In both ξ_i is the standard normal distribution N(0,1) and ε_i is the type I extreme value distribution.

These results indicate by the σ parameter the size of correlation across the choices of the same participant, over all the participants. It is thus the size of inter-participant variation, regarding the left or right alternative in the first approach, and regardless of the alternative in the second approach. The results are given for the complete dataset and the three definitions of repeated observations (table 9.19).

The results show there is hardly any effect of having repeated observations, regardless of the definition used. The additional estimated parameters σ or σ_i are close to zero, with just one case it being significantly different from zero (using the homoscedastic approach with UserTripID). Neither of the models provide a significant better fit according to the likelihood ratio test ($p>0.10$). These results indicate that there is no or very little correlation of choices across observations of participants. This is contrary to the expectations, especially for the definition of full trip, where essentially the sequence of choice leading to a certain route is modelled. This was expected to be important, as the current choice is conditional on the choices made at the previous locations. Adding correlation to the sequence of observations between participants thus does not add any extra explaining value in the model.

Table 9.19: *Models applying the error components approaches for repeated observations*

	UserID		**UserTripID**		**FullTripID**	
Variable	*Approach 1*	*Approach 2*	*Approach 1*	*Approach 2*	*Approach 1*	*Approach 2*
ASC1 (right)	−0.555	−0.577	−0.581	−0.583	−0.577	−0.577
ASC2 (straight)	0	0	0	0	0	0
ASC3 (left)	−0.778	−0.772	−0.771	−0.769	−0.814	−0.772
$\beta_{FFTT(\lambda)}$	−1.10	−1.05	−1.04	−1.04	−1.04	−1.05
$\beta_{\log D}$	−0.558	−0.574	−0.576	−0.574	−0.582	−0.574
β_{JC}	−0.224	−0.237	−0.239	−0.240	−0.237	−0.237
β_C	0.617	0.597	0.598	0.596	0.622	0.597
β_I	−0.666[a]	−0.618[a]	−0.617[a]	−0.624[a]	−0.621[a]	−0.618[a]
β_{TM}	−3.88	−3.89	−3.89	−3.93	−3.98	−3.98
β_A	1.40	1.39	1.39	1.40	1.40	1.39
λ	0.619	0.633	0.635	0.638	0.654	0.633
σ		0.007[a]		0.097		0.004[a]
σ_R	0.263[a]		0.096[a]		0.002[a]	
σ_L	0.020[a]		0.07[a]		0.442[a]	
Number of observations	2517	2517	2517	2517	2517	2517
Null log likelihood	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932
Final log likelihood	−878.282	−878.793	−878.708	−878.363	−878.329	−878.793
Likelihood ratio test	2605.300	2604.278	2604.448	2605.137	2605.207	2604.278
Pseudo ρ^2	0.597	0.597	0.597	0.597	0.597	0.597
Adjusted pseudo ρ^2	0.592	0.592	0.592	0.592	0.592	0.592
Number of parameters	12	11	12	11	12	11

[a] Non-significant estimate at a 95% level.
[b] Simulation has been performed with 1000 Halton draws

Table 9.20: *Comparing the base model with the heteroscedastic alternative model*

Variable	Base	Heteroscedastic
ASC1 (right)	−0.577	−0.561
σ_{right}		0.009[a]
ASC2 (straight)	0	0
ASC3 (left)	−0.772	−0.929
σ_{left}		1.06
$\beta_{FFTT(\lambda)}$	−1.05	−1.06
$\beta_{\log D}$	−0.574	−0.607
β_{JC}	−0.237	−0.240
β_{C}	0.597	0.777
β_{I}	−0.618[a]	−0.560[a]
β_{TM}	−3.89	−4.39
β_{A}	1.39	1.49
λ	0.633	0.672
Number of observations	2517	2517
Null log likelihood	−2180.932	−2180.932
Final log likelihood	−878.793	−877.866
Likelihood ratio test	2604.277	2606.131
Pseudo ρ^2	0.597	0.597
Adjusted pseudo ρ^2	0.592	0.592
Number of parameters	10	12

[a] Non-significant estimate at a 95% level.

Heteroscedastic logit

Disregarding the repeated observations, a model was applied assuming heteroscedasticity (different variances) in the alternatives. This indicates if the choices for the alternatives all had the same attractiveness. It is basically the same model as used in approach 1 described above, but in this case, there is no sequence of choices assumed. In this approach, the alternative with the lowest variance is used, which was again the straight ahead exit. Because there is no sequence of choices, the resulting standard deviations can be given as variation around the mean of the unobservable elements in the utility, the alternative specific constants.

The results given in table 9.20 show no improvement in model fit (likelihood ratio test has p>0.10) with the added heteroscedastic errors in two alternatives. The alternative with the lowest variance has been selected as the one normalized, where the variance is set to zero. The right and left exit thus have a variance independent of the straight exit (both independently normal distributed), of which that of the right exit is not significantly different from zero. The additional variance around the mean also has hardly any influence on the mean itself for the right exit. The left exit is a different story, where the mean has become more negative, showing less preference compared to the straight ahead exit. With a standard deviation of 1.06, this indicates a significantly different value from zero (with a variance of 1.11). This indicates that in around 18% of the observations there is a preference for the left exit (cause by the symmetry of the normal distribution). There is thus quite some taste variation for the left alternative, compared to both other alternatives, which was not found when the variation was measured for different participants (sequence of observations).

The taste variation for the left exit is something that is not constant for a participant, but varies within the participants. Allowing for taste variation in the alternative causes a few of the parameters to have quite some difference in estimated coefficients (wether the exit brings the participant closer and the effect of traffic message on jams on the exit route), while the others show almost no difference. This indicates that these parameters are not confounded with the variation with the left (and right) exit, but are very well estimated on their own. The other two coefficients are thus now expected to be closer to their real (scaled) values, as the confounding is reduced and transferred to the taste variation for the exits. The effects of the route bringing the participant closer and traffic messages of a traffic jam on the route both are increased compared to the base model.

Random parameter logit

The last models involve applying the other approach in mixed logit; random parameters. Using the base models given in table 9.13, equation (4.17) is used in three instances. The first assumes the coefficient for freeflow travel time (maintaining the Box-Cox transformation) is applied, using the normal distribution. $\beta_{FFTT}(\lambda)$ is then distributed with N(β_{FFTT},σ_{FFTT}), which is the same as $\beta_{FFTT} + \sigma_{FFTT}N(0,1)$. The second instance assumes a lognormal distribution for the coefficient of freeflow travel time. This allows for this coefficient to be strictly negative, as it is expected that travel time cannot have a positive utility. This restriction is of course violated by the normal distribution. The third instance assumes all coefficients in the base model are normally distributed. The fourth and last instance considers both the correlation of a sequence of choice by the same participant (using userID) and a normal distribution for all coefficients, which are assumed to be independent of the sequence of observations. A fifth model, assuming heteroscedasticity, together with taste variation (either for all parameters or for freeflow travel time) did not converge after 1000 iterations. Examination of the results after 1000 iterations showed the model to be unidentifiable as there is not enough information available over the combination of parameters, especially for traffic information. In all cases, the estimated σ_x is assumed to be equal for all alternatives.

The first model in table 9.21 applies taste variation for the freeflow travel time parameter. This model shows a significant improvement in model fit ($p<0.05$). The Box-Cox transformation of the freeflow travel time is maintained and estimated simultaneously. Adding taste variation for the coefficient of freeflow travel time has quite an impact on both the mean of the coefficient and the estimated transformation factor λ. The mean is decreased to -3.83 compared to -1.05 for the model without taste variation. The standard deviation of the normally distributed coefficient is estimated to be 2.39, which indicates that around 5% of all observations indicate a positive utility attributed to freeflow travel time. Even though it is just 5%, it is still unreasonable to expect participant to prefer a longer route. The λ parameter also is quite low, indicating a larger effect for longer freeflow travel times.

In order to accommodate taste variation which is limited to negative values, a lognormal distribution was applied. The lognormal distribution is related to

Table 9.21: *Various models applying mixed logit*

Variable	Base	Normal FFTT	Lognormal FFTT	Normal base	Normal panel
ASC1 (right)	0	0	0	0	0
σ_{right}					0.126[a]
ASC2 (straight)	0.577	0.625	0.717	1.20	0.668
ASC3 (left)	−0.194[a]	−0.106[a]	−0.020[a]	0.141[a]	−0.128[a]
σ_{left}					0.341
$\beta_{FFTT(\lambda)}$	−1.05	−3.83	−0.462[a]	−0.443[a]	−2.14
σ_{FFTT}		2.39	2.73	0.400[a]	0.897
$\beta_{\log D}$	−0.574	−0.644	−0.763	−2.20	−0.743
σ_D				2.09	0.849
β_{JC}	−0.237	−0.226	−0.324	−1.17	−0.219
σ_{JC}				1.13	0.218
β_C	0.597	0.445	0.547	3.42[a]	0.489
σ_C				4.45	0.338[a]
β_I	−0.618[a]	−0.736[a]	−1.05	−8.27	−1.11
σ_I				14.6	3.26
β_{TM}	−3.88	−4.57	−48.2[a]	−27.6	−19.5
σ_{TM}				4.70[a]	9.81
β_A	1.39	1.25	1.30	2.15[a]	1.52
σ_A				1.63[a]	0.543[a]
λ	0.633	0.241	0.920	2.14	0.443
Number of observations	2517	2517	2517	2517	2517
Null log likelihood	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932
Final log likelihood	−878.793	−858.104	−852.125	−832.979	−851.360
Likelihood ratio test	2604.277	2645.657	2657.613	2695.906	2659.144
Pseudo ρ^2	0.597	0.607	0.609	0.618	0.610
Adjusted pseudo ρ^2	0.592	0.601	0.604	0.605	0.601
Number of parameters	10	11	11	17	19

[a] Non-significant estimate at a 95% level.
[b] Simulation has been performed with 1000 Halton draws

the normal distribution. If X is normally distribution, $Y = \exp X$ is lognormally distributed. As the lognormal distribution is strictly positive it is possible to enter it in the utility function using the negative value of the distribution, making it strictly negative. The estimated values are the parameters of the normal distribution, which can be used to calculate the mean and standard deviation of the lognormal distribution, using respectively $m = \ln\left(\frac{\mu^2}{\sqrt{\sigma^2+\mu^2}}\right)$ and $s = \sqrt{\ln\left((\frac{\sigma}{\mu})^2 + 1\right)}$ where μ is the estimated mean and σ is the estimated standard variation given in table 9.21. The resulting distribution has a mean of -1.1128 and a standard deviation of 1.246. As a side effect of applying the lognormal distribution, the effect of the box-cox transformation is removed. As the λ value is close to one (and is not significantly different from one), the transformation almost keeps the original value of freeflow travel time intact. This is caused by the property of the lognormal distribution where if the variable X is distribution lognormally $X \sim \log\mathcal{N}(\mu, \sigma^2)$, then $\ln(X) \sim \mathcal{N}(\mu, \sigma^2)$. As the values of freeflow travel time are now multiplied by a lognormal distribution, the effect of the Box-Cox transformation cancels out. Interestingly, the use of the lognormal distribution also has effects on the parameters for traffic information and incidents on the routes. The effect of traffic information is no longer significant, even though its value. This is due to the even larger standard error of the coefficient for traffic information. The estimated coefficient for incidents is significantly different from zero in this model. Compared to the base model, the model with a lognormal distribution has a significant better model fit (Likelihood ratio test, $p<0.05$).

In the model assuming a normal distribution for all coefficients in the base model, the pseudo ρ^2 indicates an improvement in model fit, which is confirmed by the likelihood ratio test when comparing with the base model or either of the models assuming only taste variation around freeflow travel time. However, the improving model fit also causes quite some difference in the values of the coefficients. The values for the mean and standard deviation of the freeflow travel time are both now estimated to be insignificant from zero. This effect might be caused by the increase of the λ value of the Box-Cox transformation, which is little over 2. This high value of λ indicates, contrary to expectations, that an increase of a unit travel time for higher values of travel time is more important than the same increase at low values of travel time. Instead, the mean for delay is significantly increased with a value of -2.20. Comparing with the model with just a normal distribution for the coefficient of freeflow travel time, the standard deviation of delay is in the same range, making the movement to delay more reasonable. The same effect occurs with the coefficient for the amount of junctions for each route, which is also significantly higher than in the other models. The effect of an incident on the route is also greatly increased, but this coefficient also shows a very large standard deviation, which would mean that in around 28% of the choices, the route with an incident shows a positive coefficient. This is a very unlikely consequence. The effect for traffic messages also shows a higher value, but the standard deviation in this case is not significant. The effect of the advice for a certain alternative is not significant for both the mean and standard deviation. This is unexpected, as the number of times the advised exit is chosen is very high. Based on the combination of strange effects, this model is not preferred.

The next model applied the same normal distribution to all coefficients to measure taste variation, but also includes an alternative specific variance, which is correlated over observations of the same participant (based on the definition of userID). This model separates the variation between participants (captured in the taste variation around the mean of the alternative specific constant) and between observations of all participants (the other taste variations). As such, it mimics both within and between participant variation in preferences. The model fit is again an improvement compared to the base model and the model applying a normal distribution to freeflow travel time ($p<0.05$), but not for the lognormal distributed freeflow time model ($p>0.10$). The additional parameters in the estimation for the standard variations thus don't add much more explanation of the choice. However, the specific values do indicate certain effects in the choices, which makes them interesting to discuss. The estimated correlations within participants for a certain exit show that there is no significant correlation (or variance between participants) for the right exit, but there is for the left exit. Some participants, regardless of their specific trip, are more likely to select the left exit than others. The effects and sizes for all the other coefficients are as expected, whereas the allowance of taste variation causes the mean to change in its value compared to the base model. This specifically applies to the freeflow travel time, traffic messages and incidents. Traffic messages have also quite a large standard deviation, indicating a large variation in taste over observations for this parameters. The use of traffic information might thus very well depend on the specific location of the choice, which causes the variation. The same applies to incidents. For the advised alternative however, there is no significant variation. Once advice is available, in all observations the same effects are found. This is compliant with the previously mentioned percentage of choices adhering to the advised exit.

9.4 Discussion

9.4.1 Choices

In a real world situation, a driver does not have unlimited time to choose between the exits at a junction. Moreover, a driver often also needs to make a route choice at an earlier location, as a junction can have separate lanes for each exit. In the AIDA RCS, a participant only had to choose between the exits. The time left to make a choice was presented to the participant as a blue bar with decreasing size (see figure 6.5). If a participant did not make a choice in time, the AIDA RCS has to make this choice for the participant. The AIDA RCS selects the exit which is part of the fastest route at that time (section 6.3). This might not be the choice a participant would have made, but enables the participant to keep concentrated on the experiment. If the chosen exit would be random, the resulting route might seem really unreasonable to the participant, and make him or her lose interest in performing as deemed appropriate in a morning commute. During the experiment these missed choices have been recorded. Although some participants noted the choice being made for them, none of them mentioned it as problematic.

In general, most choices were made by participants themselves. This indicates the length of the links in the road network, together with the ratio of the time in reality and in the simulation were appropriate for most choices to be made in time. The amount of choices made by a user did vary over the scenarios, incidents and type of information that was available. As expected, scenario 3 with its north-south orientation showed the most problems for participants to choose in time. Dynamic route guidance clearly helped the participants in this scenario.

9.4.2 Compliance

The trips that were made with dynamic route guidance show a very high degree of compliance, ranging from 83% tot 99% based on compliance at individual junctions. These compliance rates to the advice provided to participants is high, compared to findings in previous studies. Chen & Jovanis (2003) found compliance rates of around 90% of guidance at individual junctions, which varies over factors such as occurrence of congestion and incidents and education level. In a simulated environment, Geweke & Zumkeller (2006) found compliance rates of 84% for pre-trip recommendations and 73% for en-route recommendations. In two recent studies by Lee et al. (2010); Ben-Elia et al. (2010) compliance rates were found of 94% tot 78% for weekdays and 95% to 87% for weekends or 82% for highly accurate information to 59% for information with low accuracy. In an older study, Bonsall (1992) found that 35% of all trips complete adhered to the advice for a full route, with a 72% compliance at individual junctions.

These findings show that compliance is increased over time, which might be caused by an increasing performance, accuracy and reliability of the route guidance provided by currently available navigation systems. Car drivers are increasingly familiar with the workings of a navigation system, and quite often completely trust the route advice given. In this experiment, no variations were used for the quality, reliability or accuracy of the information and advice given. Combined with the hypothetical network that was used, participants obviously showed a high degree of compliance with the advice given. This could indicate that participants trusted the advice given, even in incident situations, or that they felt uncertain about their own choices. It could also mean a participant would have made the same route choice when no advice would have been provided. However, the amount of choices for the fastest exit (which is equal to the advised exit) shows the largest improvement between textual information and dynamic route guidance. It is expected that especially the reliance on the route guidance is the most important, as is shown by the importance of this factor in route choice, as found in the post-trip questions (section 8.3).

9.4.3 Choice models

Results

In the previous section, a large range of various utility functions and model have been estimated all based on the same proposed route choice model (see section 4.3.4). The models range from the basic MNL model, with its restrictive assumptions of independent and identically distributed error terms to the mode complex mixed logit models, which allow to relax these strict assumptions. All of these models have been found possible to estimate on the data set available, although not all parameters could be used (because of not enough datapoints for certain combinations). All models have shown to have a rather high model fit based on the likelihood ratio test and the pseudo ρ^2. The two best performing models based on this pseudo ρ^2 are the MNL model with a scale parameter added for the different scenarios (a pseudo ρ^2 of 0.606) and the base model assuming normally distributed taste variations for all parameters (a pseudo ρ^2 of 0.618). Both are a little higher than the base model including the effects of incidents, traffic information and advice (a pseudo ρ^2 of 0.597).

Behaviorally, the base model including effects and the mixed logit model assuming normally distributed coefficients are of the most interest. The first because the estimated coefficients are easily interpreted and all have a clear meaning. Transferring such a model which includes different scale parameters to a real situation is very difficult, because the scale parameters depend on the combination of viable routes and origin and destination. This dependence is difficult to predict beforehand. In case of the mixed logit model, the estimated coefficients also show a variation around the mean. Such a value is also possible to interpret, as it indicates the variation of how drivers rate the importance of the attributes of the alternatives. However, this mixed logit model is much more time consuming to estimate, because of the necessary simulation. Next, the normal distribution also violates some likely restrictions on the estimated coefficients (because it has large tails, it also means some drivers will have a positive estimated coefficient for travel time for example, which is highly unlikely). Therefore, the model that is chosen to be the best is the MNL model with the additional effects.

Consecutive choices

In this research, the approach to use each decision point as a separate observation does not take a conditionality of consecutive decisions into account. Instead, the observations at each decision point in the network are used as separate data elements. The resulting model predicts the probability of selecting a certain exit at each decision point, based on the assumption that for each exit the current fastest route to the destination will be used. Combining these probabilities, a probability of observing a complete route can be calculated, which relates to the adaptive route choice model as mentioned in section 4.3. It is also possible to predict the probability of each route in the network between a certain destination and origin to be calculated using the data of these individual observations.

Apart from the reasons mentioned in section 4.3, two other reasons are provided for not estimating an adaptive route choice model or routing policy model. First, the future state of the road network is not known, apart from the duration of the accidents. The rest of the traffic state (in terms of travel times, link flows, routes of other vehicles) is not known. The same applies to the probability of these future traffic states. Without such information, it is not possible to estimate a routing policy model. Combined with the size of the road network this creates a large computational burden, as for all links and all times, the possible traffic states have to be determined. The routing policy has, to our knowledge, only been applied in small networks with only a few routes.

Second, relating to both the adaptive route and route policy models, the number of observations for each incident situation is low, as each incident is used only three times for each participant. This results in a maximum of 30 observed routes per incident. As shown in section 8.4, there is a large overlap of the routes chosen. Together with the variation of traffic information available, each combination only has 10 routes. Such an amount of data with only little variation, does not convey enough information to be able to estimate a choice model.

Another approach was used to investigate the correlation of the consecutive choices, by assuming the data to be panel data, having multiple responses from one respondent. Three different definitions for these observations have been tested, based on the individual participant, the index of the trip for each participant and the index of each individual trip. This last definition closely resembles the adaptive routing policy, as it assumes correlation for the choices within a certain trip. Interestingly, regardless of the definition used, there is no significant improvement in the choice models estimated. This indicates that based on the current choice set, the choices for an exit at a decision point are independent from previous or future decision points. In other words, participant made a choice at each junction for just the next exit, not for a complete route.

This results however might be caused by two effects. The first is the use of a route choice simulator. In such an environment, participants are forced to choose consciously at each intersection. Although in reality, drivers also have to choose at each intersection, such a choice might be made in a more routine way, especially in a well known environment. Based on previous experience, they choose a global route to their destination, which is only re-evaluated at certain junctions, or when traffic information warrants a new decision (Bovy & Stern, 1990).

A second cause for this finding is the composition of the choice set. In this research, the choice set is based on the available exits at a junction (with a maximum of three). Each exit is labelled as right, straight ahead or left. The parameters of these alternatives are based on the current fastest route starting at that exit. The choice set thus has a maximum of three labelled alternatives, whereas the alternatives in a route choice set based on complete paths are unlabelled (Bovy, 2009). The amount of alternatives considered by an individual is generally not larger than five, but the construction of this individual specific choice set is not necessarily based on just the exits at a junction. In general the individual choice set is based on previous knowledge, current information, habits, constraints, preferences, etc. It is very well possible that the choice set used in this study does not include one or all of the alternatives that existed in the choice set of each participant. As argued by (Bekhor et al., 2006; Prato & Bekhor, 2007; Bliemer & Bovy, 2008; Bovy, 2009; Schüssler, 2010), the composition of the choice set is of great importance for the estimated choice models. However, as this approach used a labelled alternative contrary to the other approaches, based on the angle of the exit to the entrance of the junction, the choice set is expected to be less relevant. Next, as participants had no previous experiences in the hypothetical network, it can be assumed they try to optimize their route choice in terms of travel time at first. Based on this assumption the parameters for the alternatives were calculated for the fastest route given the exit.

Validity

The estimated choices models showed a significant influence of various route and exit based attributes on the utility of a route, and thus on the probability of being chosen. As expected, advice for an exit is a very significant variable, regardless of the dataset that was used for estimation. This is in accordance with the previously found high rate of compliance with the route guidance. Applying the estimated models to predict the probability of choosing an exit for data not used in the estimation showed a good match compared to the observed probabilities. As such, the internal validity of the models, both in terms of the utility functions and the estimated values is rather high. This is also confirmed by the high pseudo ρ^2.

However, to investigate the external validity of the route choice model, it is necessary to have data for equal situations. Such data is not available, which makes it impossible to test the external validity of the developed models. As other route choice models concern the choice of unlabelled routes, the results of this study can only be tested for internal validity. The external validity of models based on data gathered using route choice simulators has been proven in previous studies section 6.4. Together with the internal validity of the models, the appropriate signs and sizes of the estimated coefficients, the models are expected to be valid.

9.5 Summary

The choices at each decision point were presented, analysed and modelled in this chapter. The research questions posed in section 1.2.3 are answered with the results presented in this chapter, together with the results presented in chapter 8.

Participants showed to be able to make a route choice in time in around 96% of all choices made. The time needed for participants to make a choice was thus appropriate, with the speedup factor chosen such that participants were able to make the “right” choice.. In scenario 3, most choices were missed, which is caused by the orientation of this scenario. With textual traffic information and again with dynamic route guidance, participants made more choices in time, even with the incidents on the road network in these trips. The same improvement is found for the choices made by the participant, but comparing their choice with the fastest exit at that instant. In total 88% of the choices were for the fastest exit. For trips made with dynamic route guidance which advised the fastest route, this increases to 94%. The effects of incidents is most apparent in scenario 4, where dynamic route guidance shows the largest increase in choices for the fast exit. However, for the incidents 2, 4, 5 and 6, the provision of information has a significant effect on the choice for the fastest exit. For dynamic route guidance, most participants followed the advice.

The individual choices at each junction were used to create two datasets; one which included all choices, and the other included all choices made at strategic locations in the road network. Using these datasets, various choice models were estimated. These choice models all have the same definition for the choice set and the alternatives in common. The choice set consists of a maximum of three labelled alternatives; the exit links at each junction, which were labelled according to their angle to the entry link as left, straight ahead or right. The attributes of the alternatives are based on the routes starting at the exit links, using the fastest route to the destination.

The base models used the following attributes of the alternatives in the utility function:

- alternative specific constant;
- a Box-Cox transformed freeflow travel time;
- a log transformed delay;
- the amount of intersections, and;
- dummy indicating if the end of the exit is closer to the destination (only for the complete dataset).

For the complete dataset, the straight-ahead exit has a preference over both others, whereas left is even less preferred than right. The strategic dataset shows no differing preferences for either of the exits. For both datasets, the coefficient for freeflow travel time has the largest size, indicating the largest effect on the choice for an exit. However, as the Box-Cox transformation is used, with a value of below 1, the additional effect of an extra unit of freeflow travel time at larger values is smaller than at small values of freeflow travel time. Next, delay is important in the choice between the alternatives. Participants thus try to minimize their freeflow travel time, but also their total travel time. The complexity of a route, defined as the number of decision points in the route, is also significant, but in case of the total dataset, not as important as the fact that the route brings the participant closer to his/her destination.

Adding attributes relating to the traffic situation and traffic information resulted in a significant improvement of the models, in terms of model fit using the likelihood ratio test and increasing values of pseudo ρ^2. For the complete dataset, having an advice for an exit is more important than the freeflow travel time of this route. Having received a message on congestion on that route is the most important, but this is most likely due to the small amount of choices for such a route. Incidents on the route are not significant, but do provide a sound behavioral meaning as such a route has an increase in its disutility. The strategic dataset showed a slightly different result, as having received advice for an exit is not as important as the freeflow travel time for this exit.

Further investigations showed that the correction factors that deal with the overlap in routes which violates the IID principle of the MNL model, give no significant improvement. This can be explained from the location of the overlap and the construction of the choice set using labelled alternatives. Other additions which account for heterogeneity and heteroscedasticity concern taste and choice variations between participants and trips. The results indicated that there is significant variation between scenarios, especially for scenario 3. It also showed that there is no significant learning effect, as the correlation for the choices made by one participant is not significant. It also showed the assumption that a choice at a junction is independent of the choices at previous junctions.

The models have been used to test for internal validity, using a part of the dataset for estimation. The other choices are used to predict. In all cases, the prediction is good, with only small differences between the predicted and observed probabilities for an alternative. The model estimated on scenario 3 has the largest differences with that respect, which is as expected for this scenario. In general, the base models have a very good model fit based on the pseudo ρ^2 and together with the high internal validity, the exit choice models are expected to provide a good prediction of exit choices, together with a sound valuation of the attributes that are important in choosing a route in accident situations.

If you are worried about falling off the bike,
you'd never get on.

— Lance Armstrong

Chapter 10

Conclusions and discussion

In this final chapter, an overview of the results presented in this thesis is provided. The first section (section 10.1) presents an overview and the main findings with respect to the research objective and research questions, including the chosen approach, followed by a short summary of the conclusions of this research in section 10.2. Section 10.3 gives the implications of these results, followed by a reflection on the findings and the approach used in section 10.4. The chapter finishes with recommendations for future and further research in section 10.5.

10.1 Overview of results

The main research objective of this research was *to study the impact of traffic information on route choice behaviour in non-recurrent traffic situations.* This objective consists of three elements which were further researched in this thesis:

- non-recurrent traffic situations;
- traffic information;
- route choice behaviour.

To achieve this objective first a framework to describe non-recurrent traffic situations was developed. Second, an internet survey was conducted to investigate the need and preferences for traffic information in non-recurrent traffic situations. Third, a route choice simulator was developed which was used to conduct a study on the route choice behaviour in non-recurrent traffic situations with traffic information. The main objective was divided in three main research questions in section 1.2. This section recalls the research questions and summarises the findings.

10.1.1 Non-recurrent traffic situations

1. *Which traffic situations are non-recurrent?*

Traffic flows causing congestion are labelled as non-recurrent if the cause and occurrence of congestion can be defined as transient; it only exists for a specific period of time at a specific location. This is the opposite of recurring congestion, which is persistent or regular; it occurs every day. However, traffic flows consist of individual travellers, each with different experiences and expectations. A definition of non-recurrence for traffic situations solemnly focussing on congestion and its characteristics ignores the differences of the travellers in this congestion.

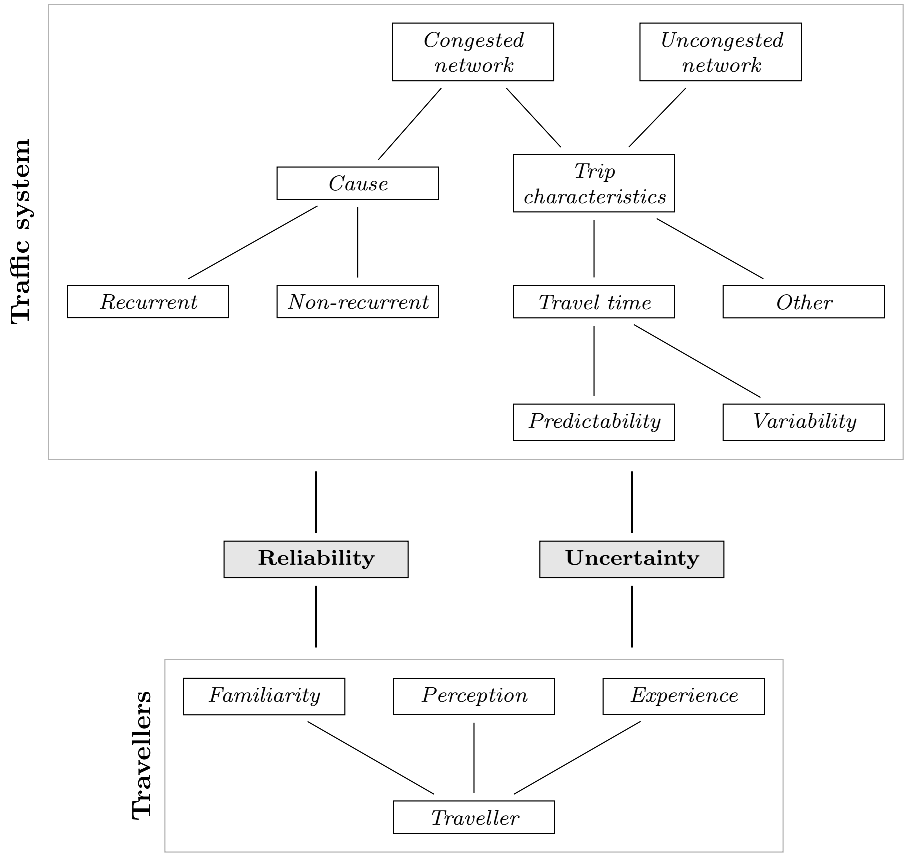


Figure 10.1: *Framework for non-recurrent traffic situations*

To this end, a framework was developed based on the findings from a literature survey. The framework is repeated in figure 10.1.

In this framework, the concept of a traffic situation is used, extending that of congestion or a traffic pattern. A traffic situation is defined as the traffic pattern or congestion a traveller encounters in his current trip. This explicitly introduces the traveller into a definition for non-recurrent traffic situations.

A traveller has a varying degree of experience with both the road network (its physical lay-out) and the traffic patterns that occur on this road network. The experience with the road network concerns the well known routes, which relates to spatial knowledge. The experience with the traffic patterns concerns the full range of temporal and spatial knowledge, ranging from the experience with speeds for specific lanes at a specific bottleneck to the seasonal and daily variation of traffic flows on the complete network. For a trip it is possible to distinguish between the (external) information, the specific experience and the global experience, applied to both the road network and the traffic patterns. With increasing trip frequency, the more specific experience is available and the less information is needed or acquired.

In all cases, the experience always concerns a perceived traffic pattern and road network, which means that each traveller will have a unique perception of the road network and traffic patterns. What is regular and recurrent to one traveller might be non-recurrent to another, depending on the experience. To what extent a traffic situation is non-recurrent thus depends on the traveller for which the traffic situation is described.

The traffic patterns which will often be related to non-recurrent traffic situations are the non-recurrent congestions such as accidents and road works. These traffic patterns cause a variability in and unpredictability of the travel time (for a specific group of travellers). The characteristics of a traveller that describe the degree of non-recurrence of a traffic situation are the familiarity and experience with the road network and traffic patterns together with the perception hereof.

Reliability and certainty relate the subjective elements (as considered by the traveller) of the traffic situation to the objective elements (as described by using the (un)congested road network and its cause and effects). Both sides have their own dynamics, but meet each other in the middle by a certain degree of reliability and certainty.

Three example causes for a change in traffic patterns are discussed. These situations are an accident, road works and a large event. For different travellers, these situations are discussed, as not all drivers will have the same experience or familiarity. A driver with multiple trips with the same road works will build an experience for the traffic patterns that will occur for these road works and will have a different degree of non-recurrence for this traffic situation compared to a driver who has build no experience with the resulting traffic patterns.

10.1.2 Traffic information

2. *What traffic information do drivers prefer in non-recurrent situations?*

Non-recurrent congestion is often mentioned as the most appropriate situation for which ATIS can be used to inform travellers of the circumstances and provide them with alternatives. However, the framework developed for non-recurrent traffic situations indicates that even in non-recurrent congestion, not all travellers experience the same degree of non-recurrence. The need and preference for traffic information in such situations was expected to differ between drivers, based on their personal characteristics.

An internet survey was conducted to investigate the preferences of drivers for different non-recurrent traffic situations. A conceptual model was developed, based on an extensive literature review, to describe the relations between traffic situations and personal characteristics on one end and the attributes, types and contents of information on the other end. The survey was designed using the conceptual model and the framework for non-recurrent traffic situations.

The results of the survey are based on the answers of 598 Dutch respondents that completed the survey, of which a large part was recruited as representation of the population of Dutch car drivers. The first part of the survey concerned their personal characteristics and the attitude towards mobility of the participants as a general classification regardless of the common socio economic characteristics such as age and level of education.

The second part of the survey concerned the current usage of traffic information. Based on the findings in literature, travellers that are already using traffic information are also more inclined to use traffic information. Most respondents already use traffic information before and during travel, and they indicate to be satisfied with this information. The most used source of traffic information before and during the trip is the radio, which means only congestion information is used.

The third part of the survey investigated the relationship between information contents and traffic situations. It showed that the traffic situation of a traveller influences the choice of contents of traffic information, together with other characteristics of the trip such as the trip purpose. Regardless of the type of situation, the most preferred content for traffic information is advice for the fastest route, followed by extended traffic information (location, length, cause and expected duration of traffic jams on the complete network) and expected time of arrival. For specific situations, other types of information are chosen more often, such as advice for a route with the best scenery, but never as much as the top three.

The preference for traffic information also differs over travellers. Three main descriptives of a traveller are of influence, regardless of the situation. These are the attitude towards mobility, age and usage of traffic information during travel. Travellers currently using information preferred more often the extended traffic information whereas older travellers prefer advice for the shortest route more. The attitude towards mobility describes the way a traveller looks at mobility. For instance, an "enjoyer" sees being mobile as the ability to be social and enjoy travelling. "Competitors" on the other hand see mobility as a symbol of independence and like to be "better" than others. As was expected, the different preferences for traffic information relates to the attitude towards mobility.

The last part of the survey concerned the attributes of traffic information in one traffic situation, and was designed as a stated preference survey to be used in a discrete choice model. Five alternatives were presented to the respondents of which four had three attributes with different levels. These attributes were the cost, timeliness and reliability of the traffic information. The fifth alternative was the no information option.

The results show that 60% of the respondents were willing to pay for traffic information, regardless of the type or other attributes. Within this 60%, most participants choose advice or personal advice, which confirms the finding for the contents of traffic information. Of the attributes of traffic information, costs are the most important, followed by reliability and timeliness. This order of attributes is the same, regardless of the underlying model assumptions used to estimate the choice models. The specific preference for a type of traffic information and attributes did differ between the different attitudes for mobility. For example, "competitors" have the lowest rating of costs, which indicates they are more willing to pay than others.

10.1.3 Route choice behaviour

3. *What are the effects of traffic information on route choice during non-recurrent traffic situations?*

An experiment with a route choice simulator was undertaken to investigate route choice behaviour. A route choice simulator was chosen based on the need to acquire data on route choice in non-recurrent traffic situations while the traveller receives traffic information. A route choice simulator has several advantages over other methods such as surveys or observations, such as the ability to represent realistic situations and controlling for all variations in the traffic patterns of the traveller. The AIDA Route Choice Simulator was designed as client-server architecture where the server connects to Paramics for the traffic patterns and a database for storing the results. The AIDA RCS is capable of online studies and allows multiple participants at the same time.

In total 30 participants took part in the experiment with the AIDA RCS. The participants were invited amongst friends and colleagues which means the result are not a true representation of the dutch population of car drivers. The experiment consisted of 8 trips for each participant with a particular combination of origin and destination, accidents and traffic information. For the experiment a hypothetical road network was designed, which assured that all participants would have the same starting point for experience and familiarity. The order of the trips started with no information and no accidents, followed by adding accidents and then textual information followed by dynamic route guidance. This meant that all participants had the same possibility to acquire experience.

The results for the experiment consist of the general results, the results for the individual choices and the results for the face validity of the experiment and AIDA RCS. First, the face validity was investigated using post-trip questions on the general experience and whether the participant would make the same choice in a similar real life situations. Both questions indicate a high level of validity of the AIDA RCS and the experiment as most participant said the trips were realistic or very realistic and would make similar route choices outside of the AIDA RCS.

The routes that were chosen depend on the traffic information participant received. Graphical overviews showed difference between the three kinds of information. Especially dynamic route guidance reduces the number of different routes chosen by the participants, which means that traffic information and dynamic route guidance cause a concentration of route choices. The route choice did vary between the scenarios and within the scenarios between the accidents. Especially in case an accident is located on the route that participants seemed to prefer, the accident caused participants to reroute and find a way around the accident or its consequences. With respect to travel times, accidents cause an increase of almost 10% when no information was available, corresponding to a 30% increase of delay. Traffic information causes a reduction of the average travel time for the same accidents with over 13% and dynamic route guidance adds another 6% reduction. Especially delay is reduced, which indicates participants were able to find routes that had less variability in travel time. Traffic information and route guidance had a significant impact on the travel time and delay. With respect to the individual choices, participants were able to select the fastest exit at a junction in 88% of the choice when receiving traffic information. Dynamic route guidance increased this to 94%, which meant that in almost all cases, the participants followed the advice provided.

The individual choices were also used to estimate a route choice model. The route choice model assumes that a driver selects an alternative at each decision point in the road network. The alternatives available are the exits at that decision point. In this research we assumed a maximum of three alternatives (excluding the u-turn), which are left, straight ahead and right. The alternative is considered to lead to the destination with the current fastest route, for which the important attributes are freeflow travel time, delay, number of decision points or junctions and a dummy indicating if the alternative brings the driver closer to the destination. The choice for an exit is independent of the previous choices a traveller made.

The proposed model is estimated using the multinomial logit model (MNL), assuming that the driver will maximize the utility (or minimize the disutility) of his choice by selecting the best alternative. The estimated model showed that participants preferred the straight ahead exit, followed by right and then left. Freeflow travel time is the most important attribute of the alternative, regardless of the specific exit, followed by delay. Extending the model by including attributes for accidents and traffic information and route guidance improve the ability of the model to reproduce the observed choices, with an increase in the model fit. Especially having an advice for an exit shows to be very important in the choice for this alternative. However, this depends on the data used for the estimation, as a model estimated on choice made at strategic decisions showed that freeflow travel time is more important. The relative values and order of importance of the different attributes of the routes are expected to be applicable to other situations.

The model was tested for the important assumption of independent and identically distributed errors, which is violated in the proposed model using MNL, as we assumed independence of consecutive choices. Tests showed that the model did not improve in model fit, nor did the valuation of the attributes of the alternatives change. Overlap in the alternatives also violates the assumption of the underlying MNL model (as the alternatives are no longer independent). Based on a theoretical derivation, a model including a correction factor should outperform models without such a factor. Two different correction factors (Path Size and Path Size Correction) have been applied. However, neither of these factors provided a better model fit compared to the basic MNL model, which is most likely due to the composition of the choice set, as only exits are used as alternatives. This approach differs from other route choice models where a choice set does not necessarily consider all exits at a decision point, but uses an algorithm to calculate the alternatives, which means overlap in the routes exists in both the beginning and end of a route. Our approach only contains overlap in the end of a route, and essentially considers a choice for the next link, not for the complete route.

The models have been tested for internal validity by using a part of the dataset for estimation and the remaining part for prediction. In all cases, the predictions show only small difference with the aggregated observed choices. Based on this internal validity and the very good model fit, the proposed model is expected to provide a good prediction of choices at junctions.

10.2 Conclusions

After presenting the findings of this research, the main conclusions are presented. This is done for the three elements mentioned in the objective: non-recurrent traffic situations, traffic information and route choice behaviour.

Non-recurrent traffic situations are an important factor in the need and preference for traffic information
It was found that the different non-recurring traffic situations show a significant impact on the preference of travellers for the contents of traffic information. A traveller will prefer an advice in a trip for which he has no previous experiences, whereas for a trip where a large set of experiences is available, the preference is towards information which allows the traveller to make better choices himself.

The framework for non-recurrent traffic situations is a useful way to describe the need and preference for traffic information
The framework described shows that apart from the traffic patterns and the congestion that occurs, also the individual within these traffic patterns is important. As all travellers have different experiences and perceptions, even for the same traffic patterns, the need for traffic information differs between travellers, as well as the preference. This was confirmed in the survey for preferences where traffic situation was an important factor in the preferred contents.

The preference for traffic information depends on personal characteristics
It was found that personal characteristics are causing differences in the preferred contents of traffic information. Especially the fact that a traveller currently uses traffic information, their age and their attitude towards mobility cause differences.

Drivers are willing to pay for traffic information
Regardless of the type of traffic information, in 60% of the cases a traveller chose to receive and pay for traffic information, while the option to have no traffic information was available. This clearly indicates a willingness to pay for traffic information.

Traffic information leads to travel time savings for the individual traveller in case of accidents
When drivers were provided with textual traffic information, which also included information on the location, severity and expected duration of an accident, they were able to reduce their travel time. A combination of textual information and dynamic route guidance even further reduced the travel time. The reduction of travel time is to a large extent caused by a reduction in delay.

Traffic information leads to a concentration of travellers on the same route
Traffic information, especially route guidance, indicated a reduction in the number of routes chosen by drivers. The currently fastest route is advised to all drivers with the same origin and destination, which automatically leads to a concentration of these drivers on the same route.

Drivers make a route choice at each junction
The proposed model assumed that drivers choose an exit at each junction. The model has been estimated and tested for its underlying assumptions, which showed

that accounting for the violations (such as correlated alternatives because of overlap and dependence of consecutive choices) introduced did not improve the model fit, while retaining the relative weights of the attributes in the model. This leads us to conclude that the model could be used to predict route choices in similar situations.

Freeflow travel time and advice are the important attributes of routes
Within the choice model, the assumption was made that the choice for an exit concerned that exit with the fastest route to the destination. Of the attributes of each alternative, the freeflow travel time and advice for an exit were found to be the most important in determining the choice.

10.3 Implications

The conclusions and overview of results have several implications for policy makers, consultants, road operators and service providers. A summary of the important implications will be presented here.

- The consideration of different preferences even for the same traveller, depending on the traffic situation and other characteristics such as trip purpose indicate a need to personalize the traffic information towards drivers. This involves a delicate process in asking travellers for their personal and trip characteristics, as well as accurately describing experiences with traffic situations.
- As drivers are willing to pay for traffic information, it should be possible to derive a valid business case for accurate and personalized traffic information. This relates to the previous implication, as it considers the personalization of traffic information as a service.
- Investing in detecting accidents in particular and traffic patterns in general in urban areas as this allows for a better provision of traffic information on these traffic situations. The availability of such traffic information allows travellers to reduce their travel time significantly in case of accidents on their preferred routes.
- With the increasing availability and use of navigation systems, the choice for a route has become a choice for the following link. This choice is largely based on the advice provided by the navigation system. Improving the underlying algorithms for the advice, based on the preferred choices by the participant makes the system even more credible, even in recurrent traffic situations, which in turn could change the habitual behaviour of travellers.

10.4 Reflections

This research was based on the idea that the need for traffic information (and consequently, the preference and use of traffic information) is highest in trips with a high level non-recurrence. However, the current traffic information is not particularly targeted at such trips. A regular trip has different implications for the traffic information a traveller needs, compared to a trip trough an unknown road network, or a trip in which unpredictable events cause significant variation in traffic patterns. As a consequence, a driver will vary in his route choices in these situations. The reflection on the research presented here considers this idea and the approach used in this research.

The framework developed to describe non-recurrent traffic situations was based on the available insights. It extends and relates the important concepts in the classification of trips which are already described in the current literature. The framework combined the spatial cognition and cognition of traffic patterns of an individual traveller with the notion of congestion, its causes and effects. The framework does not provide a definitive answer in how travellers classify traffic situations, but allows a researcher to describe traffic situations in terms which are well understood.

The preference for traffic information was investigated using an internet survey, in which respondents could indicate the information they would prefer to receive in a specific situation. The preference was indicated without actually using the information, in a stated preference survey. This means it is not possible to relate the stated preference for traffic information to the actual usage of traffic information in these situations. Nevertheless, the fact that preferences differed between traffic situations and between different drivers underpin the idea of this research that traffic information should not be treated as a constant service.

The route choice simulator allowed us to investigate the effects of traffic information and accidents on the route choice of drivers. A route choice simulator combines the advantages of a stated preference and revealed preference research, but nonetheless remains a virtual environment. As the consequences of a choice are difficult to represent in a virtual environment, a driver might make different choices in a real life situations. However, the validity of similar environment for acquisition of choice behaviour have been proven to provide valid results. The environment and the experiment were also found to be realistic by the participants of this research. This allows us to investigate the route choice behaviour, but we advise to interpret the specific values with caution, especially for the transferability to other situations. Based on the face validity and validity of similar environments, we expect the relative valuation of the elements in the choice model as well as the general results will be applicable in real life situations.

The model proposed for the route choices assumed the driver to be maximizing utility. This assumption is often criticised to be invalid. However, the model is treated as a way of indicating the important aspects in route choice behaviour in non-recurrent traffic situations. The model was developed to acquire information on the (relative) importance of various attributes of a route such as travel time, occurrence of accidents and the availability of an advice. Next to the investigation of choice behaviour, we think the approach used here, by assuming drivers choice at each junction, is a concept which is very well applicable in real life situations.

10.5 Further research

This section discusses some directions for further research, based on the issues, assumptions and findings in this research.

First, the AIDA RCS is used to study a specific traffic situations involving a commute and accidents with drivers without experience, as we used a hypothetical road network. We recommend to extend the application of the AIDA RCS to other traffic situations, also involving road works, large events and others as it would allow to further study the relationship between non-recurrent traffic situations and the need and use of traffic information with respect to route choice behaviour. We also recommend to differ between hypothetical road networks to study the concepts and control for the experience of drivers, and existing road networks, which allow for an even more realistic behaviour.

Second, we recommend to include other personal characteristics into the framework for non-recurrent traffic situations, especially learning behaviour. Learning behaviour is found to be an important aspect in decision making for travel, and also concerns the learning of road networks, traffic patterns and the accuracy and quality of traffic information. We excluded this behaviour in this study to reduce the amount of variables, which does not mean that learning behaviour is not important.

Third, we recommend to extend the investigation of driver preferences to real life situations which travellers have actually experienced. This would enable to combine the preference for and use of traffic information in different trips.

Fourth, the AIDA RCS developed for this research is a flexible and extendible environment to study route choice behaviour. In this research we focussed on route choice in non-recurrent traffic situations, but many other elements in route choice behaviour can be studied. For example, studying the adherence to an advice for the most economical, safest or environmentally friendliest route. Another possibility is to expand the number of simultaneous users to investigate interactions, perhaps together with road user charging and slot management (providing a few alternative time windows and a routes for a traveller, together with varying costs from which a traveller chooses; the slot).

Finally, we recommend to explore the exit choice model. Similar approaches are already described, but do not explicitly include non-recurrent traffic situations. Further studies using the exit choice model would have to include revealed choice data, which would allow for validation of the findings presented here. Other directions for research are the reproduction of complete routes based on the choices at individual junctions, to investigate the ability of the exit choice model to reproduce a complete route as a sequence of individual choices. Expanding the choice model to include possible future states of the network is also recommended.

Bibliography

Abdalla, M.F. & Abdel-Aty, M.A. (2006), Modeling Travel Time Under ATIS Using Mixed Linear Models, *Transportation*, vol. 33, no. 1, pp. 63–82.

Abdel-Aty, M.A. & Abdalla, M.F. (2004), Understanding Drivers Route Choice Under Long-Term Pretrip and Short-Term En-Route Traffic Information Using Generalized Estimating Equations, *Journal of Transportation Engineering*, vol. 130, no. 6, p. 777.

Abdel-Aty, M.A. & Abdalla, M.F. (2006), Examination of Multiple Mode/Route-Choice Paradigms Under ATIS, *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 3, pp. 332–348.

Abdel-Aty, M.A., Kitamura, R. & Jovanis, P.P. (1997), Using stated preference data for studying the effect of advanced traffic information on drivers' route choice, *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 1, pp. 39–50.

Adler, J.L. (2001), Investigating the learning effects of route guidance and traffic advisories on route choice behavior, *Transportation Research Part C: Emerging Technologies*, vol. 9, no. 1, pp. 1–14.

Adler, J.L. & Blue, V.J. (1998), Toward the design of intelligent traveler information systems, *Transportation Research Part C: Emerging Technologies*, vol. 6, no. 3, pp. 157–172.

Adler, J.L. & McNally, M.G. (1994), In-laboratory experiments to investigate driver behavior under advanced traveler information systems, *Transportation Research Part C: Emerging Technologies*, vol. 2, no. 3, pp. 149–164.

Adler, J.L., Recker, W.W. & McNally, M.G. (1993), A conflict model and interactive simulator (FASTCARS) for predicting enroute driver behavior in response to real-time traffic condition information, *Transportation*, vol. 20, no. 2, pp. 83–106.

AGV & ITS (1996), Gebruikerswensen verkeersinformatie; Eindrapport UNEVA-project, Technical Report 1-0663/GP/1578, Ministerie van Verkeer en Waterstaat - Adviesdienst Verkeer en Vervoer.

Allen, R.W., Ziedman, D., Rosenthal, T.J., Stein, A.C., Torres, J.F. & Halati, A. (1991), Laboratory assessment of driver route diversion in response to in-vehicle navigation and motorist information systems, *Transportation Research Record*, vol. 1306, pp. 82–91.

Allenby, G.M. & Ginter, J.L. (1995), The effects of in-store displays and feature advertising on consideration sets, *International Journal of Research in Marketing*, vol. 12, no. 1, pp. 67–80.

Anable, J. (2005), 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying travel behaviour segments using attitude theory, *Transport Policy*, vol. 12, no. 1, pp. 65–78.

Arnott, R., de Palma, A. & Lindsey, R. (1991), Does providing information to drivers reduce traffic congestion?, *Transportation Research Part A: General*, vol. 25, no. 5, pp. 309–318.

Avery, R., Burghout, W. & Andreasson, I. (2008), An Interactive Tool for Collecting Traveler Behavior Information, in *Transportation Research Board 87th Annual Meeting*, 08-0373, pp. 1–14, Washington D.C.

Avineri, E. & Prashker, J.N. (2005), Sensitivity to travel time variability: Travelers learning perspective, *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 2, pp. 157–183.

Avineri, E. & Prashker, J.N. (2006), The Impact of Travel Time Information on Travelers Learning under Uncertainty, *Transportation*, vol. 33, no. 4, pp. 393–408.

Azalient (2007), Ceejazz Client, http://www.azalient.net.

Barfield, W., Haselkorn, M., Spyridakis, J. & Conquest, L. (1991), Integrating commuter information needs in the design of a motorist information system, *Transportation Research Part A: General*, vol. 25, no. 2-3, pp. 71–78.

Bates, J., Polak, J.W., Jones, P. & Cook, A. (2001), The valuation of reliability for personal travel, *Transportation Research Part E: Logistics and Transportation Review*, vol. 37, no. 2-3, pp. 191–229.

Bazzan, A.L.C., Wahle, J. & Klügl, F. (1999), Agents in traffic modelling-from reactive to social behaviour, in *KI-99: Advances in Artificial Intelligence*, vol. 1701, edited by W. Burgard, A.B. Cremers & T. Cristaller, pp. 303–306, Springer, ISBN 978-3-540-66495-6.

Beaton, P., Chen, C. & Meghdir, H. (1998), Stated choice: a study in predictive validity using an aggregate truth set, *Transportation*, vol. 25, no. 1, pp. 55–75.

Bekhor, S., Ben-Akiva, M. & Ramming, M.S. (2006), Evaluation of choice set generation algorithms for route choice models, *Annals of Operations Research*, vol. 144, no. 1, pp. 235–247.

Bekhor, S., Ben-Akiva, M. & Scott Ramming, M. (2002), Adaptation of Logit Kernel to Route Choice Situation, *Transportation Research Record*, vol. 1805, no. 1, pp. 78–85.

Bekhor, S. & Prashker, J. (2001), Stochastic User Equilibrium Formulation for Generalized Nested Logit Model, *Transportation Research Record*, vol. 1752, no. 1, pp. 84–90.

Ben-Akiva, M. & Bierlaire, M. (2003), Discrete Choice Methods and Their Applications to Short-Term Travel Decisions, in *Handbook of Transportation Science*, edited by R.W. Hall, 1985, Chapter 2, pp. 7–37, Springer US, ISBN 978-0-306-48058-4.

Ben-Akiva, M. & Lerman, S.R. (1985), *Discrete choice analysis: Theory and application to travel demand*, MIT Press Series in Transportation Studies, MIT Press, Cambridge, Massachusetts, fifth edit edn., ISBN 0262022176.

Ben-Akiva, M., Palma, A.D. & Kaysi, I. (1991), Dynamic network models and driver information systems, *Transportation Research Part A: General*, vol. 25, no. 5, pp. 251–266.

Ben-Akiva, M. & Ramming, M.S. (1998), Lecture notes: Discrete choice models of traveler behavior in networks.

Ben-Elia, E., Erev, I. & Shiftan, Y. (2007), The combined effect of information and experience on drivers route-choice behavior, *Transportation*, vol. 35, no. 2, pp. 165–177.

Ben-Elia, E., Pace, R.D. & Bifulco, G.N. (2010), Modelling the impacts of ATIS accuracy on travellers ' route-choice behaviour and risk perception, in *European Transport Conference 2010*.

Ben-Elia, E. & Shiftan, Y. (2010), Which road do I take? A learning-based model of route-choice behavior with real-time information, *Transportation Research Part A: Policy and Practice*, vol. 44, no. 4, pp. 249–264.

van Berkum, E.C. & van der Mede, P.H. (1993), *The impact of traffic information*, Phd, Delft University of Technology.

Bhat, C.R. (1995), A heteroscedastic extreme value model of intercity travel mode choice, *Transportation Research Part B: Methodological*, vol. 29, no. 6, pp. 471–483.

Bie, J., van Arem, B. & van Berkum, E.C. (2011), Dynamic Route Choice and the Influence of a Safety Incentive: Experimental Results with Dynamic Route Choice Simulator, in *Transportation Research Board 90th Annual Meeting*, Washington D.C.

Bierlaire, M. (2003), BIOGEME: a free package for the estimation of discrete choice models, in *Swiss Transport Research Conference*, Ascone, Switzerland.

Bierlaire, M. (2008), An introduction to BIOGEME, biogeme.epfl.ch, École Polytechnique Fédérale de Lausanne.

Bierlaire, M., Chen, J. & Newman, J. (2010), Modeling Route Choice Behavior From Smartphone GPS data, Technical report.

Bierlaire, M. & Frejinger, E. (2008), Route choice modeling with network-free data, *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 2, pp. 187–198.

Bliemer, M.C.J. & Bovy, P.H.L. (2008), Impact of Route Choice Set on Route Choice Probabilities, *Transportation Research Record*, vol. 2076, pp. 10–19.

Bogers, E.A. (2009), *Traffic Information and Learning in Day-to-Day Route Choice*, Phd, Delft University of Technology.

Bogers, E.A., Bierlaire, M. & Hoogendoorn, S.P. (2007), Modeling Learning in Route Choice, *Transportation Research Record*, vol. 2014, pp. 1–8.

Bogers, E.A., Viti, F. & Hoogendoorn, S.P. (2005), Joint modeling of advanced travel information service, habit, and learning impacts on route choice by laboratory simulator experiments, *Transportation Research Record*, vol. 1926, pp. 189–197.

Bonsall, P.W. (1992), The influence of route guidance advice on route choice in urban networks, *Transportation*, vol. 19, no. 1, pp. 1–23.

Bonsall, P.W. (2004a), Route choice simulators, in *Human behaviour and traffic networks*, edited by M. Schreckenberg & R. Selten, pp. 167–191, Springer Verlag, ISBN 3540212205.

Bonsall, P.W. (2004b), Traveller Behavior: Decision-Making in an Unpredictable World, *Journal of Intelligent Transportation Systems*, vol. 8, no. 1, pp. 45–60.

Bonsall, P.W., Firmin, P.E., Anderson, M., Balmforth, P. & Palmer, I. (1997), Validating the results of a route choice simulator, *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 6, pp. 371–387.

Bonsall, P.W. & Palmer, I. (2004), Modelling drivers' car parking behaviour using data from a travel choice simulator, *Transportation Research Part C: Emerging Technologies*, vol. 12, no. 5, pp. 321–347.

Bonsall, P.W. & Parry, T. (1991), Using an interactive route-choice simulator to investigate driver's compliance with route guidance advice, *Transportation Research Record*, vol. 1306, pp. 59–68.

Bovy, P.H.L. (2007), Modeling route choice sets in transportation networks: a preliminary synthesis, in *Tristan V Conference, Phuket*, January, pp. 1–21.

Bovy, P.H.L. (2009), On Modelling Route Choice Sets in Transportation Networks: A Synthesis, *Transport Reviews*, vol. 29, no. 1, pp. 43–68.

Bovy, P.H.L., Bekhor, S. & Prato, C.G. (2008), The Factor of Revisited Path Size: Alternative Derivation, *Transportation Research Record*, vol. 2076, pp. 132–140.

Bovy, P.H.L. & Stern, E. (1990), *Route choice : wayfinding in transport networks*, Kluwer Academic Publishers, Dordrecht, The Netherlands, ISBN 0792308123.

Box, G. & Cox, D. (1964), An analysis of transformations, *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 26, no. 2, pp. 211–252.

Boyd, J.H. & Mellman, R.E. (1980), The effect of fuel economy standards on the U.S. automotive market: An hedonic demand analysis, *Transportation Research Part A: General*, vol. 14, no. 5-6, pp. 367–378.

van den Broeke, A., van der Horst, R.S. & Schotanus, L. (2004), Reizigersinformatie van idee tot product, in *Colloquium Vervoersplanologisch Speurwerk 2004*, Zeist, The Netherlands.

Brookhuis, K.A., Dicke, M., Rademaker, R., Grunnekemeijer, J., Nijenhuis, M., Benedictus, N. & Eser, E. (2008), The effects of travel information presentation on driver behaviour, in *Human Factors for assistance and automation*, edited by D. de Waard, F. Flemisch, B. Lorenz, H. Oberheid & K. Brookhuis, pp. 431 – 444, Shaker Publishing.

Buijs, A. (2008), *Statistiek om mee te werken*, Noordhoff Uitgevers B.V., 8th editio edn., ISBN 9020731866.

Burnett, G. (1998), *'Turn Right at the King's Head': Drivers' Requirements for Route Guidance Information*, Phd, Loughborough University.

Cardell, N.S. & Dunbar, F.C. (1980), Measuring the societal impacts of automobile downsizing, *Transportation Research Part A: General*, vol. 14, no. 5-6, pp. 423–434.

Carson, R.T. & Groves, T. (2007), Incentive and informational properties of preference questions, *Environmental and Resource Economics*, vol. 37, no. 1, pp. 181–210.

Carson, R.T., Louviere, J.J., Anderson, D.a., Arabie, P., Bunch, D.S., Hensher, D.A., Johnson, R.M., Kuhfeld, W.F., Steinberg, D., Swait, J.D., Timmermans, H.J. & Wiley, J.B. (1994), Experimental analysis of choice, *Marketing Letters*, vol. 5, no. 4, pp. 351–367.

Cascetta, E., Nuzzolo, A., Russo, F. & Vitetta, A. (1996), A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, in *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, pp. 697–711, Lyon, France, Elsevier Science.

Cascetta, E. & Papola, A. (2001), Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand, *Transportation Research Part C: Emerging Technologies*, vol. 9, no. 4, pp. 249–263.

CBS (2011), Statline, http://statline.cbs.nl/.

Chen, P.S. & Mahmassani, H.S. (1993), Dynamic interactive simulator for studying commuter behavior under real-time traffic information supply strategies, *Transportation Research Record*, vol. 1413, pp. 12–21.

Chen, P.S.T. & Mahmassani, H.S. (1999), Dynamics of Urban Commuter Behaviour Under Real-Time Traffic Information.

Chen, P.S.T., Srinivasan, K.K. & Mahmassani, H.S. (1999), Effect of Information Quality on Compliance Behavior of Commuters Under Real-Time Traffic Information, *Transportation Research Record*, vol. 1676, pp. 53–60.

Chen, W. & Jovanis, P.P. (2003), Analysis of driver en-route guidance compliance and driver learning with ATIS using a travel simulation experiment, in *Transportation Research Board 81th Annual Meeting*, pp. 1–22, Washington D.C.

Cheng, S. & Long, J. (2007), Testing for IIA in the Multinomial Logit Model, *Sociological Methods & Research*, vol. 35, no. 4, pp. 583–600.

Chorus, C.G. (2007), *Traveler response to information*, Phd, Delft University of Technology.

Chorus, C.G., Arentze, T.A. & Timmermans, H.J. (2009), Traveler compliance with advice: A Bayesian utilitarian perspective, *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 3, pp. 486–500.

Chorus, C.G., Arentze, T.A., Timmermans, H.J., Molin, E.J. & Van Wee, B. (2007a), Travelers' Need for Information in Traffic and Transit: Results from a Web Survey, *Journal of Intelligent Transportation Systems*, vol. 11, no. 2, pp. 57–67.

Chorus, C.G., Molin, E.J., Arentze, T.A., Hoogendoorn, S.P., Timmermans, H.J. & Van Wee, B. (2007b), Validation of a multimodal travel simulator with travel information provision, *Transportation Research Part C: Emerging Technologies*, vol. 15, no. 3, pp. 191–207.

Chorus, C.G., Molin, E.J. & Van Wee, B. (2006a), Travel choice adaptation through information provision: insights from a literature review, in *Transportation Research Board 86th Annual Meeting*, January 2007.

Chorus, C.G., Molin, E.J. & Van Wee, B. (2006b), Use and Effects of Advanced Traveller Information Services (ATIS): A Review of the Literature, *Transport Reviews*, vol. 26, no. 2, pp. 127–149.

Chu, C. (1989), A paired combinatorial logit model for travel demand analysis, in *Proceedings of the 5th World Conference on Transportation Research*, pp. 295–309, Ventura, USA.

Cohn, N. (2009), Real-Time Traffic Information and Navigation, *Transportation Research Record*, vol. 2129, no. -1, pp. 129–135.

Conquest, L., Spyridakis, J., Haselkorn, M. & Barfield, W. (1993), The effect of motorist information on commuter behavior: Classification of drivers into commuter groups, *Transportation Research Part C: Emerging Technologies*, vol. 1, no. 2, pp. 183–201.

Daly, A. & Zachary, S. (1978), Improved multiple choice models, in *Determinants of Travel Choice*, vol. 335, edited by D. Hensher & M. Dalvi, p. 357, Teakfield, Farnborough, England, Sussex.

Dia, H. (2002), An agent-based approach to modelling driver route choice behaviour under the influence of real-time information, *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 5-6, pp. 331–349.

Dia, H. & Panwai, S. (2007), Modelling drivers' compliance and route choice behaviour in response to travel information, *Nonlinear Dynamics*, vol. 49, no. 4, pp. 493–509.

Dicke, M., Bouwers, F.L., Koenders, E., Kok, R., van der Veer, R. & Brookhuis, K.A. (2004), Human Factors in the design of travel and traffic information, in *Human Factors in Design*, edited by D. de Waard, K.A. Brookhuis & C. Weikert, 2004, pp. 95–107, Shaker Publishing, Maastricht, The Netherlands.

Dicke, M. & Brookhuis, K.A. (2005), Klantvriendelijke reisinformatie, in *Verkeerskunde DVM Congres*, pp. 6–9.

van Dijck, T. (2007), Discrete choice models; dynamic route choice modelling under traffic information, Technical report, University of Twente, Enschede, The Netherlands.

Emmerink, R.H., Axhausen, K.W., Nijkamp, P. & Rietveld, P. (1995), The potential of information provision in a simulated road transport network with non-recurrent congestion, *Transportation Research Part C: Emerging Technologies*, vol. 3, no. 5, pp. 293–309.

Enrique Fernández L., J., de Cea Ch, J. & Germán Valverde, G. (2009), Effect of advanced traveler information systems and road pricing in a network with non-recurrent congestion, *Transportation Research Part A: Policy and Practice*, vol. 43, no. 5, pp. 481–499.

Ettema, D. & Timmermans, H.J. (2006), Costs of travel time uncertainty and benefits of travel time information: Conceptual model and numerical examples, *Transportation Research Part C: Emerging Technologies*, vol. 14, no. 5, pp. 335–350.

Firmin, P.E. (1995), *The use of simulated travel environments to investigate driver route choice behaviour*, Phd, University of Leeds.

Fosgerau, M., Frejinger, E. & Karlström, A. (2009), Route choice modeling without route choice, in *European Transport Conference 2009*, pp. 1–14.

Frejinger, E. (2007), Random sampling of alternatives in a route choice context, *Proceedings of the European Transport Conference*.

Frejinger, E. (2008), *Route choice analysis: data, models, algorithms and applications*, Phd, École Polytechnique Fédérale de Lausanne.

Frejinger, E. & Bierlaire, M. (2007), Capturing correlation with subnetworks in route choice models, *Transportation Research Part B: Methodological*, vol. 41, no. 3, pp. 363–378.

Frejinger, E., Bierlaire, M. & Ben-Akiva, M. (2009), Sampling of alternatives for route choice modeling, *Transportation Research Part B: Methodological*, vol. 43, no. 10, pp. 984–994.

Fujii, S. & Gärling, T. (2003), Application of attitude theory for improved predictive accuracy of stated preference methods in travel demand analysis, *Transportation Research Part A: Policy and Practice*, vol. 37, no. 4, pp. 389–402.

Gale, N., Golledge, R.G., Pellegrino, J. & Doherty, S. (1990), The acquisition and integration of route knowledge in an unfamiliar neighborhood, *Journal of Environmental Psychology*, vol. 10, no. 1, pp. 3–25.

Gao, S. & Chabini, I. (2006), Optimal routing policy problems in stochastic time-dependent networks, *Transportation Research Part B: Methodological*, vol. 40, no. 2, pp. 93–122.

Gao, S., Frejinger, E. & Ben-Akiva, M. (2008), Adaptive Route Choice Models in Stochastic Time-Dependent Networks, *Transportation Research Record*, vol. 2085, pp. 136–143.

Gärling, T. & Axhausen, K.W. (2003), Introduction: Habitual travel choice, *Transportation*, vol. 30, pp. 1–11.

Gentile, G. & Papola, A. (2006), An alternative approach to route choice simulation: the sequential models, in *European Transport Conference 2006*, Strasbourg.

Geweke, S. & Zumkeller, D. (2006), A simulation game to explore the effects of information on traveler behavior, in *Transportation Research Board 85th Annual Meeting*, Washington D.C.

Golledge, R.G. (2002), Dynamics and ITS: Behavioural responses to information available from ATIS, in *In perpetual Motion: Travel behaviour research opportunities and application challenges*, edited by H.S. Mahmassani, pp. 81–126, Pergamon, Amsterdam, NL.

Hall, R.W. (1993), Non-recurrent congestion: How big is the problem? Are traveler information systems the solution?, *Transportation Research Part C: Emerging Technologies*, vol. 1, no. 1, pp. 89–103.

Hall, R.W. (1996), Route choice and advanced traveler information systems on a capacitated and dynamic network, *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 5, pp. 289–306.

Hallenbeck, M.E., Ishimaru, J.M. & Nee, J. (2003), Measurement of Recurring versus Non-Recurring Congestion, Technical report, Washington State Transportation Center, Seattle, Washington.

Hart, P., Nilsson, N. & Raphael, B. (1968), A Formal Basis for the Heuristic Determination of Minimum Cost Paths, *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107.

Hato, E., Taniguchi, M., Sugie, Y., Kuwahara, M. & Morita, H. (1999), Incorporating an information acquisition process into a route choice model with multiple information sources, *Transportation Research Part C: Emerging Technologies*, vol. 7, no. 2-3, pp. 109–129.

Hensher, D.A. (1997), A practical approach to identifying the market potential for high speed rail: A case study in the sydney-canberra corridor, *Transportation Research Part A: Policy and Practice*, vol. 31, no. 6, pp. 431–446.

Hensher, D.A. & Greene, W.H. (2002a), Specification and estimation of the nested logit model: alternative normalisations, *Transportation Research Part B: Methodological*, vol. 36, no. 1, pp. 1–17.

Hensher, D.A. & Greene, W.H. (2002b), The Mixed Logit Model: The State of Practice.

Hensher, D.A. & Greene, W.H. (2003), The mixed logit model: the state of practice, *Transportation*, vol. 30, no. 2, pp. 133–176.

Hensher, D.A., Rose, J.M. & Greene, W.H. (2005), *Applied choice analysis: a primer*, Cambridge University Press, New York, USA, ISBN 0521844266.

Hess, S. (2005), *Advanced discrete choice models with applications to transport demand*, Phd, Imperial College London.

Hess, S. & Rose, J. (2006), Some lessons for working with repeated choice data, Technical Report 0, ETH, Eidgenössische Technische Hochschule Zürich, Institut für Verkehrsplanung und Transportsysteme, Zürich.

Hess, S., Rose, J.M. & Hensher, D.A. (2008), Asymmetric preference formation in willingness to pay estimates in discrete choice models, *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, no. 5, pp. 847–863.

Hilbers, H., van Eck, J. & Snellen, D. (2004), Behalve de dagelijkse files: over betrouwbaarheid van reistijd, Technical report, Ruimtelijk Planbureau, Den Haag, The Netherlands.

Hoogendoorn-Lanser, S. (2005), *Modelling travel behaviour in multi-modal networks*, Phd, Delft University of Technology.

Iida, Y., Akiyama, T. & Uchida, T. (1992), Experimental analysis of dynamic route choice behavior, *Transportation Research Part B: Methodological*, vol. 26, no. 1, pp. 17–32.

Jou, R.C., Lam, S.H., Liu, Y.H. & Chen, K.H. (2005), Route switching behavior on freeways with the provision of different types of real-time traffic information, *Transportation Research Part A: Policy and Practice*, vol. 39, no. 5, pp. 445–461.

Kantowitz, B.H., Becker, C.A. & Barlow, S.T. (1993), Assessing Driver Acceptance of IVHS Components, in *Human Factors and Ergonomics Society Annual Meeting Proceedings*, pp. 1062–1066, Human Factors and Ergonomics Society.

Karl, C.A. (2003), *The Learning Driver: Meeting Traveller Information Needs*, Ph.D. thesis, The Australian Graduate School of Entrepreneurship.

Karl, C.A. & Bechervaise, N. (2003), The driver as an active learner: customising real time traveller information, in *10th International Conference on Travel Behaviour Research*, pp. 10–15, Lucerne, Switzerland.

Katsikopoulos, K.V., Duse-Anthony, Y., Fisher, D.L. & Duffy, S.a. (2000), The Framing of Drivers' Route Choices when Travel Time Information Is Provided under Varying Degrees of Cognitive Load, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 42, no. 3, pp. 470–481.

Katsikopoulos, K.V., Duse-Anthony, Y., Fisher, D.L. & Duffy, S.A. (2002), Risk attitude reversals in drivers' route choice when range of travel time information is provided., *Human factors*, vol. 44, no. 3, pp. 466–73.

Katteler, H. & Broeders, W. (2002), Weggebruikers en on-line verkeersinformatie in de auto; Gebruikersonderzoek RIC, UNEVA-2, Technical report, Ministerie van Verkeer en Waterstaat - Adviesdienst Verkeer en Vervoer.

Kennisinstituut voor Mobiliteitsbeleid (2008), File-facts, Technical report.

Kennisinstituut voor Mobiliteitsbeleid (2010), Mobiliteitsbalans 2010, Technical report.

Kenyon, S. & Lyons, G. (2003), The value of integrated multimodal traveller information and its potential contribution to modal change, *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 6, no. 1, pp. 1–21.

Khattak, A.J., Pan, X., Williams, B., Rouphail, N. & Fan, Y. (2008), Traveler Information Delivery Mechanisms: Impact on Consumer Behavior, *Transportation Research Record*, vol. 2069, pp. 77–84.

Khattak, A.J., Schofer, J.L. & Koppelman, F.S. (1993), Commuters' enroute diversion and return decisions: Analysis and implications for advanced traveler information systems, *Transportation Research Part A: Policy and Practice*, vol. 27, no. 2, pp. 101–111.

Khattak, A.J., Yim, Y. & Prokopy, L.S. (2003), Willingness to pay for travel information, *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 2, pp. 137–159.

Kim, K. & Vandebona, U. (1999), User Requirements and Willingness To Pay for Traffic Information Systems: Case Study of Sydney, Australia, *Transportation Research Record*, vol. 1694, no. 1, pp. 42–47.

Klügl, F. & Bazzan, A.L.C. (2002), Simulation of adaptive agents: learning heuristics for route choice in a commuter scenario, in *Proceedings of the first international joint conference on Autonomous agents and multiagent systems part 1 - AAMAS '02*, p. 217, ACM Press, New York, New York, USA, ISBN 1581134800.

Koopmans, C. & Kroes, E. (2004), Estimation of congestion costs in the Netherlands, Technical Report 28, SEO, Amsterdam, The Netherlands.

Koppelman, F.S. & Wen, C.H. (2000), The paired combinatorial logit model: properties, estimation and application, *Transportation Research Part B: Methodological*, vol. 34, no. 2, pp. 75–89.

Koutsopoulos, H.N., Lotan, T. & Yang, Q. (1994), A driving simulator and its application for modeling route choice in the presence of information, *Transportation Research Part C: Emerging Technologies*, vol. 2, no. 2, pp. 91–107.

Koutsopoulos, H.N., Polydoropoulou, A. & Ben-Akiva, M. (1995), Travel simulators for data collection on driver behavior in the presence of information, *Transportation Research Part C: Emerging Technologies*, vol. 3, no. 3, pp. 143–159.

Kraan, M., Mahmassani, H.S. & Huynh, N. (2000), Traveler Responses to Advanced Traveler Information Systems for Shopping Trips: Interactive Survey Approach, *Transportation Research Record*, vol. 1725, pp. 116–123.

Lappin, J. (2000), Advanced Traveler Information Service (ATIS): Who are ATIS Customers?, Technical report, John A. Volpe National Transportation Systems Center, Cambridge, Massachusetts.

Lappin, J. & Bottom, J. (2001), Understanding and predicting traveler response to information: a literature review, Technical report, John A. Volpe National Transportation Systems Center, Cambridge, Massachusetts.

Lee, S., Son, H., Lee, S. & Choi, I. (2010), The analysis of drivers' compliance rate in car navigation systems, in *World Conference on Transport Research*, pp. 1–14, Lisbon, Portugal.

Levinson, D. (2003), The value of advanced traveler information systems for route choice, *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 1, pp. 75–87.

Li, H., Guensler, R. & Ogle, J. (2005), Analysis of Morning Commute Route Choice Patterns Using Global Positioning System-Based Vehicle Activity Data, *Transportation Research Record*, vol. 1926, no. 1, pp. 162–170.

van Lint, J., van Zuylen, H.J. & Tu, H. (2008), Travel time unreliability on freeways: Why measures based on variance tell only half the story, *Transportation Research Part A: Policy and Practice*, vol. 42, no. 1, pp. 258–277.

Lomax, T., Schrank, D., Turner, S. & Margiotta, R. (2003), Selecting travel reliability measures, Technical Report May 2003, Texas Transportation Institute.

Lotan, T. (1997), Effects of familiarity on route choice behavior in the presence of information, *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 3-4, pp. 225–243.

Lotan, T. & Koutsopoulos, H.N. (1993), Models for route choice behavior in the presence of information using concepts from fuzzy set theory and approximate reasoning, *Transportation*, vol. 20, no. 2, pp. 129–155.

Louviere, J.J., Hensher, D.A. & Swait, J.D. (2000), *Stated choice methods: analysis and application*, Cambridge University Press, ISBN 9780521788304.

Lu, X., Gao, S. & Ben-Elia, E. (2011), Information impacts on route choice and learning behavior in a congested network : An experimental approach, in *Transportation Research Board 90th Annual Meeting*, pp. 1–18, Washington D.C.

Lyons, G. (2001), From Advanced Towards Effective Traveller Information Systems, in *Travel Behaviour Research: The Leading Edge*, edited by D.A. Hensher, pp. 813–826, International Association of Travel Behaviour Research, Pergamon, ISBN 0 08 043924 1.

Lyons, G. (2006), The role of information in decision-making with regard to travel, *Intelligent Transport Systems, IEE Proceedings*, vol. 153, no. 3, pp. 199–212.

Mahmassani, H.S. & Herman, R. (1990), Interactive Experiments for the Study of Trip maker Behaviour Dynamics in Congested Commuting Systems, in *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*, edited by P. Jones, pp. 272–298, Gower, Aldershot, England.

Mahmassani, H.S., Huynh, N.N., Srinivasan, K.K. & Kraan, M. (2003), Tripmaker choice behavior for shopping trips under real-time information: model formulation and results of stated-preference internet-based interactive experiments, *Journal of Retailing and Consumer Services*, vol. 10, no. 6, pp. 311–321.

Mahmassani, H.S. & Jayakrishnan, R. (1991), System performance and user response under real-time information in a congested traffic corridor, *Transportation Research Part A: General*, vol. 25, no. 5, pp. 293–307.

Mahmassani, H.S. & Jou, R.C. (2000), Transferring insights into commuter behavior dynamics from laboratory experiments to field surveys, *Transportation Research Part A: Policy and Practice*, vol. 34, no. 4, pp. 243–260.

Mahmassani, H.S. & Liu, Y.H. (1999), Dynamics of commuting decision behaviour under advanced traveller information systems, *Transportation Research Part C: Emerging Technologies*, vol. 7, no. 2-3, pp. 91–107.

Mahmassani, H.S. & Stephan, D. (1988), Experimental investigation of route and departure time choice dynamics of urban commuters, *Transportation Research Record*, vol. 1203, pp. 69–84.

Mannering, F., Kim, S.G., Ng, L. & Barfield, W. (1995), Travelers' preferences for in-vehicle information systems: An exploratory analysis, *Transportation Research Part C: Emerging Technologies*, vol. 3, no. 6, pp. 339–351.

May, A.D. (1990), *Traffic flow fundamentals*, Prentice Hall, ISBN 0139260722.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behaviour, in *Frontiers in Econometrics*, edited by P. Zarembka, pp. 105–142, Academic Press, New York, USA.

McFadden, D. (1978), Modeling the choice of residential location, in *Spatial Interaction Theory and Planning Models*, edited by A. Karlqvist, L. Lundqvist, F. Snickars & J. Weibull, pp. 75–96, Elsevier Science Ltd., Amsterdam.

McFadden, D. & Train, K.E. (2000), Mixed MNL models for discrete response, *Journal of Applied Econometrics*, vol. 15, no. 5, pp. 447–470.

van der Mede, P.H. & van Berkum, E.C. (1991), Modeling route choice, inertia and responses to variable message signs, *6th International Conference on Travel Behaviour Research*.

Mehndiratta, S.R., Kemp, M., Lappin, J. & Brand, D. (1999), What advanced traveler information system information do users want? Evidence from in-vehicle navigation device users, *Transportation Research Record*, vol. 1679, pp. 41–49.

Mensonides, W. (2004), *User preferences towards traffic information in case of road incidents*, Msc., University of Twente.

Might, M. (2011), The Illustrated Guide to a Ph.D.

Ministerie van Verkeer en Waterstaat (2003), Gebruikersbehoeften Reisinformatie, Technical Report April, Ministerie van Verkeer en Waterstaat, Adviesdienst Verkeer en Vervoer.

Ministerie van Verkeer en Waterstaat (2004), Reisinformatie en weggebruiker, Technical report, Adviesdienst Verkeer en Vervoer, Rotterdam, The Netherlands.

Ministerie van Verkeer en Waterstaat (2005), *Nota mobiliteit: naar een betrouwbare en voorspelbare bereikbaarheid*, september, Ministerie van Verkeer en Waterstaat.

Ministerie van Verkeer en Waterstaat (2008), Beleidskader Benutten, Technical report.

Ministerie van Verkeer en Waterstaat (2010), Kwartaalmonitor bereikbaarheidsontwikkeling Hoofdwegennet, Technical report.

Munizagaa, M.A., Heydecker, B.G. & Ortúzar, J.d.D. (2000), Representation of heteroskedasticity in discrete choice models, *Transportation Research Part B: Methodological*, vol. 34, no. 3, pp. 219–240.

Nee, J., Hallenbeck, M.E. & Legg, B. (2001), Designing Arterial Traveler Information Systems from a Travelers Perspective, in *ITS world congress*, 206.

Need, Y. (2002), Is mobiliteit voor iedereen fun? Verschillende belevingen van mobiliteit, Technical Report september, Adviesdienst Verkeer en Vervoer.

Ng, L., Barfield, W. & Mannering, F. (1995), A survey-based methodology to determine information requirements for advanced traveler information systems, *Transportation Research Part C: Emerging Technologies*, vol. 3, no. 2, pp. 113–127.

Noland, R. & Polak, J.W. (2002), Travel time variability: A review of theoretical and empirical issues, *Transport Reviews*, vol. 22, no. 1, pp. 39–54.

Ozbay, K., Datta, A. & Kachroo, P. (2001), Modeling Route Choice Behavior with Stochastic Learning Automata, *Transportation Research Record*, vol. 1752, no. 1, pp. 38–46.

Palma, A.D. & Picard, N. (2005), Route choice decision under travel time uncertainty, *Transportation Research Part A: Policy and Practice*, vol. 39, no. 4, pp. 295–324.

Palmer, I., Bonsall, P.W. & Shires, J.D. (1998), Driver response to guidance: a comparison of data from a route-choice simulator and a self-completion SP exercise, in *European Transport Conference 1998*.

Papinski, D. & Scott, D.M. (2011), A GIS-based toolkit for route choice analysis, *Journal of Transport Geography*, vol. 19, no. 3, pp. 434–442.

Parkany, E., Gallagher, R. & Viveiros, P. (2004), Are attitudes important in travel choice?, *Transportation Research Record*, vol. 1894, pp. 127–139.

Peeta, S. & Yu, J.W. (2005), A Hybrid Model for Driver Route Choice Incorporating En-Route Attributes and Real-Time Information Effects, *Networks and Spatial Economics*, vol. 5, no. 1, pp. 21–40.

Peirce, S. & Lappin, J. (2004), Why Don't More People Use Advanced Traveler Information?: Evidence from the Seattle Area, in *Transportation Research Board 83rd Annual Meeting*, Washington D.C.

Penttinen, M., Anttila, V. & Sandberg, H. (2003), Its-services: User needs and transport policy goals, in *10th World Congress on ITS*, pp. 1–7.

Petrella, M. & Lappin, J. (2004), Los Angeles and Seattle: a comparative analysis of customer response to online traffic information, in *Transportation Research Board 83rd Annual Meeting*, pp. 1–11, Washington D.C.

Polak, J.W. & Jones, P. (1993), The acquisition of pre-trip information: A stated preference approach, *Transportation*, vol. 20, no. 2, pp. 179–198.

Polydoropoulou, A., Ben-Akiva, M., Khattak, A.J. & Lauprete, G. (1996), Modeling Revealed and Stated En-Route Travel Response to Advanced Traveler Information Systems, *Transportation Research Record*, vol. 1537, no. 1, pp. 38–45.

Polydoropoulou, A., Gopinath, D.A. & Ben-Akiva, M. (1997), Willingness To Pay for Advanced Traveler Information Systems: SmarTraveler Case Study, *Transportation Research Record*, vol. 1588, no. 1, pp. 1–9.

Polydoropoulou, A. & Tsirimpa, A. (2010), The role of Flexibility on Traffic Information Acquisition, in *World Conference on Transport Research*, pp. 1–12, Lisbon, Portugal.

Prashker, J. & Bekhor, S. (1998), Investigation of Stochastic Network Loading Procedures, *Transportation Research Record*, vol. 1645, no. 1, pp. 94–102.

Prato, C.G. (2004), *Latent factors and route choice behaviour*, Phd, Politecnico di Torino.

Prato, C.G. (2009), Route choice modeling: past, present and future research directions, *Journal of Choice Modelling*, vol. 2, no. 1, pp. 65–100.

Prato, C.G. & Bekhor, S. (2007), Modeling Route Choice Behavior: How Relevant Is the Composition of Choice Set?, *Transportation Research Record*, vol. 2003, no. 1, pp. 64–73.

Prato, C.G., Bekhor, S. & Pronello, C. (2011), Latent variables and route choice behavior, *Transportation*.

Quadstone (2007), Paramics Online, http://www.paramics-online.com.

Quattrone, A. & Vitetta, A. (2011), Random and fuzzy utility models for road route choice, *Transportation Research Part E: Logistics and Transportation Review*, vol. 47, no. 6, pp. 1126–1139.

Raadsen, M.P. (2005), *Design and Implementation of a Dynamic Route choice Simulator*, Msc, University of Twente.

Raadsen, M.P. & Muizelaar, T.J. (2005), Ontwikkeling van een dynamische routekeuze simulator, in *Colloquium Vervoersplanologisch Speurwerk 2005*, Antwerpen, Belgium.

Ramming, M.S. (2001), *Network knowledge and route choice*, Phd, Massachusetts Institute of Technology.

Redmond, L. & Mokhtarian, P.L. (2001), The positive utility of the commute: modeling ideal commute time and relative desired commute amount, *Transportation*, vol. 28, no. 2, pp. 179–205.

Ridwan, M. (2003), Fuzzy preference based route choice model, in *Transportation Research Board 82nd Annual Meeting*, vol. 616, Washington D.C.

Schofer, J.L. & Khattak, A.J. (1993), Behavioral issues in the design and evaluation of advanced traveler information systems, *Transportation Research Part C: Emerging Technologies*, vol. 1, no. 2, pp. 107–117.

Schofer, J.L., Koppelman, F.S. & Charlton, W. (1997), Perspectives on Driver Preferences for Dynamic Route Guidance Systems, *Transportation Research Record*, vol. 1588, no. 1, pp. 26–31.

Schüssler, N. (2010), *Accounting for similarities between alternatives in discrete choice models based on high-resolution observations of transport behaviour*, Phd, Eidgenössische Technische Hochschule ETH Zürich.

Snelder, M., Immers, B. & Wilmink, I. (2004), De begrippen betrouwbaarheid en robuustheid nader verklaard, in *Colloquium Vervoersplanologisch Speurwerk*.

Sparmann, J. (1991), Benefits of Dynamic Route Guidance Systems as Part of a Future Oriented City Traffic Management System, in *Proceedings of VNIS '91 - Vehicle Navigation and Information Systems Conference*, pp. 839–847.

Srinivasan, K.K. & Mahmassani, H.S. (2000), Modeling Inertia and Compliance Mechanisms in Route Choice Behavior Under Real-Time Information, *Transportation Research Record*, vol. 1725, pp. 45–53.

Srinivasan, K.K. & Mahmassani, H.S. (2003), Analyzing heterogeneity and unobserved structural effects in route-switching behavior under ATIS: a dynamic kernel logit formulation, *Transportation Research Part B: Methodological*, vol. 37, no. 9, pp. 793–814.

Stern, E. (2004), Spatio-temporal patterns of subjectively reported congestion in Tel Aviv metropolitan area, *Journal of Transport Geography*, vol. 12, no. 1, pp. 63–71.

Stopher, P. (2004), Reducing road congestion: a reality check, *Transport Policy*, vol. 11, no. 2, pp. 117–131.

Streff, F. & Wallace, R. (1993), Analysis of driver's information preferences and use in automobile travel: Implications for Advanced Traveler Information Systems, in *Proceedings of VNIS '93 - Vehicle Navigation and Information Systems Conference*, pp. 414–418, IEEE, Ottawa, Ont. , Canada, ISBN 0-7803-1235-X.

Suhonen, N. (2007), Normative and Descriptive Theories of Decision Making under Risk: A Short Review.

Thomas, T. & Tutert, B. (2009), Een nieuwe kijk op routekeuze? Bevindingen uit het kentekenonderzoek Enschede, in *Colloquium Vervoersplanologisch Speurwerk*.

Toledo, T. & Beinhaker, R. (2006), Evaluation of the Potential Benefits of Advanced Traveler Information Systems, *Journal of Intelligent Transportation Systems*, vol. 10, no. 4, pp. 173–183.

Train, K.E. (2009), *Discrete choice methods with simulation*, Cambridge University Press, 2nd edition edn., ISBN 978-0-521-76655-5.

Tsirimpa, A., Polydoropoulou, A. & Antoniou, C. (2007), Development of a Mixed Multi-Nomial Logit Model to Capture the Impact of Information Systems on Travelers' Switching Behavior, *Journal of Intelligent Transportation Systems*, vol. 11, no. 2, pp. 79–89.

Tsirimpa, A., Polydoropoulou, A. & Antoniou, C. (2010), Development of a Latent Variable Model to Capture the Impact of Risk Aversion on Travelers Switching Behavior, *Journal of Choice Modelling*, vol. 3, no. 1, pp. 127–148.

Tu, H. (2008), *Monitoring travel time reliability of freeways*, Phd, Delft University of Technology.

Van de Kaa, E. (2010), Prospect Theory and Choice Behaviour Strategies: Review and Synthesis of Concepts from Social and Transport sciences, *EJTIR*, vol. 10, no. 10, p. 4.

Vaughn, K.M., Abdel-Aty, M.A., Kitamura, R., Jovanis, P.P. & Yang, H. (1993), Experimental Analysis and Modeling of Sequential Route Choice Under ATIS in a Simple Traffic Network, Technical report, Institute of Transportation Studies, UC Berkeley, Berkely.

Vaughn, K.M., Prasuna, R., Abdel-Aty, M.A., Kitamura, R. & Jovanis, P.P. (1995), Route choice and information use: initial results from simulation experiments, *Transportation Research Record*, vol. 1516, pp. 61–69.

Vovsha, P. (1997), Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area, *Transportation Research Record*, vol. 1607, no. 1, pp. 6–15.

Walker, J.L. (2001), *Extended Discrete Choice Models : Integrated Framework , Flexible Error Structures , and Latent Variables*, Phd, Massachusetts Institute of Technology.

Walker, J.L., Ben-Akiva, M. & Bolduc, D. (2007), Identification of parameters in normal error component logit-mixture (NECLM) models, *Journal of Applied Econometrics*, vol. 22, no. 6, pp. 1095–1125.

Wallace, R. & Streff, F. (1993), Traveler information in support of driver's diversion decisions: A survey of driver's preferences, in *Proceedings of VNIS '93 - Vehicle Navigation and Information Systems Conference*, pp. 242–246, IEEE, Ottawa, Ont. , Canada, ISBN 0-7803-1235-X.

Weijermars, W. (2007), *Analysis of urban traffic patterns using clustering*, Phd, University of Twente.

Weisser, K.I. & Horowitz, A. (2002), Perspectives and Expectations of Drivers : A Literature and Best Practices Scan, Technical Report 0092, University of Wisconsin, Milwaukee.

Wen, C.H. & Koppelman, F.S. (2001), The generalized nested logit model, *Transportation Research Part B: Methodological*, vol. 35, no. 7, pp. 627–641.

Wikipedia (2011a), Usability.

Wikipedia (2011b), User interface design.

Williams, H. (1977), On the formation of travel demand models and economic evaluation measures of user benefit, *Environment and Planning A*, vol. 9, no. 3, pp. 285–344.

Wilmink, I., van den Broeke, A. & Berghout, L. (2003), De taakverdeling tussen netwerkstructuur en verkeersmanagement, in *Colloquium Vervoersplanologisch Speurwerk*.

Wolinetz, L., Khattak, A.J. & Yim, Y. (2001), Why Will Some Individuals Pay for Travel Information When It Can Be Free? Analysis of a Bay Area Traveler Survey, *Transportation Research Record*, vol. 1759, no. 1, pp. 9–18.

Yáñez, M., Cherchi, E., Heydecker, B.G. & Ortúzar, J.d.D. (2010), On the treatment of repeated observations in panel data: efficiency of mixed logit parameter estimates, *Networks and Spatial Economics*, pp. 1–26.

Yang, H., Kitamura, R., Jovanis, P.P., Vaughn, K.M. & Abdel-Aty, M.A. (1993), Exploration of route choice behavior with advanced traveler information using neural network concepts, *Transportation*, vol. 20, no. 2, pp. 199–223.

Zhang, L. & Levinson, D. (2008), Determinants of Route Choice and Value of Traveler Information: A Field Experiment, *Transportation Research Record*, vol. 2086, pp. 81–92.

Appendix A

The survey

Behoefte aan verkeersinformatie

Verkeersinformatie is tegenwoordig een gemeengoed. Iedereen kent wel de radio boodschappen na het nieuws waarin de locatie en lengte van de files worden beschreven. Op bepaalde plaatsen wordt ook door middel van grote borden aangegeven waar de files staan of wat de reistijd is naar een bepaalde lokatie. Navigatiesystemen zoals TOMTOM kennen de mogelijkheid om bij het uitrekenen van de optimale route rekening te houden met locaties van files. Tenslotte is ook op internet uitgebreide file-informatie te vinden. Zo geeft de ANWB op haar website de locatie en lengte van alle files weer.

Deze voorbeelden maken duidelijk is dat er voor de automobilisten veel mogelijk is. De vraag is echter wat de automobilist precies wil. Welke informatie vindt de automobilist belangrijk en welke niet? Welke kenmerken van informatie zijn van belang hierbij? In welke situaties is de meeste behoefte aan welke informatie? En natuurlijk, wat doet de automobilist met de informatie die aan hem/haar gegeven wordt? Verandert zijn of haar gedrag? Worden er andere routes gekozen? Deze enquête gaat in op een aantal van deze vragen. De antwoorden worden gebruikt in een promotie onderzoek dat wordt uitgevoerd op de Universiteit Twente.

Doel
Het doel van deze enquête is het analyseren van behoeften en voorkeuren van automobilisten ten aanzien van verkeers en reisinformatie. In deze enquête wordt allereerst een aantal vragen gesteld om te bepalen wat voor type reiziger u bent. Vervolgens kunt u aangeven welke informatie u in welke situatie wenst te ontvangen. Tot slot wordt u gevraagd voor één situatie aan te geven welke soort informatie u wenst te ontvangen, rekening houdend met bijvoorbeeld de kosten van de informatie.

Instructie
De enquête bestaat uit een aantal pagina's met vragen. Lees de vragen en antwoorden alstublieft zorgvuldig door. Om op de volgende pagina te komen drukt u op de 'Verder' knop van de vragenlijst. Om terug te gaan naar een vorige pagina drukt u op de 'Terug' of 'Back' knop van uw internetbrowser. U moet dan wel de pagina herladen, met behulp van refresh. Het invullen van de enquête duurt ongeveer 15-20 minuten. Eventuele opmerkingen over de enquête kunt u aan het eind kwijt. Er zijn geen goede of foute antwoorden: het gaat om úw mening als automobilist. De door u verstrekte gegevens zullen anoniem worden verwerkt.

Verloting
Win een cadeaubon van € 20 met het invullen van deze enquête! Aan het eind van de vragenlijst kunt u aangeven of u mee wilt doen aan deze verloting.

Bij voorbaat hartelijk dank voor uw medewerking!

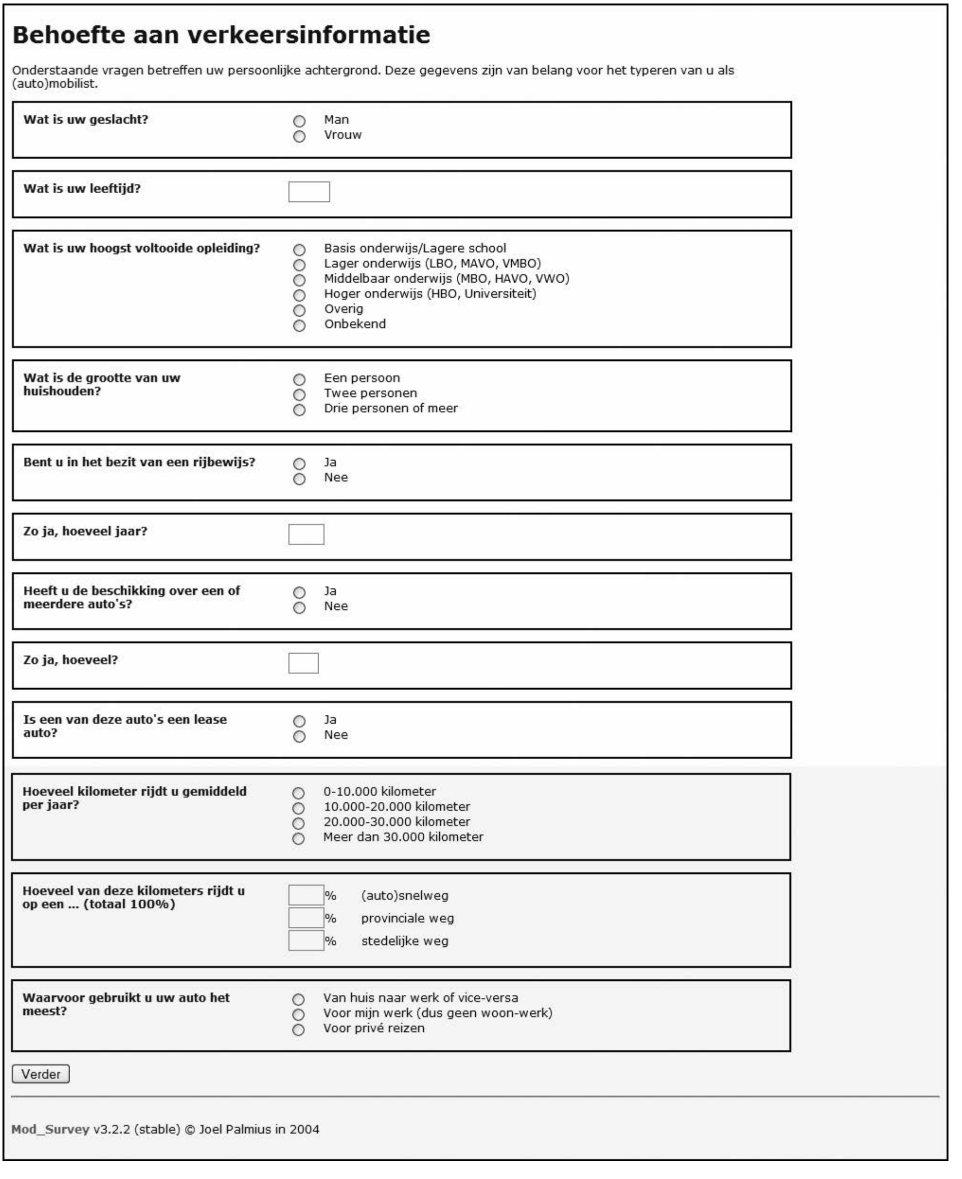
Thijs Muizelaar
Bart van Arem

Universiteit Twente
Kenniscentrum Applications of Integrated Driver Assistance (AIDA)

Door naar de enquête

Mod_Survey v3.2.2 (stable) © Joel Palmius in 2004

Figure A.1: *survey.survey*



Behoefte aan verkeersinformatie

Onderstaande vragen betreffen uw persoonlijke achtergrond. Deze gegevens zijn van belang voor het typeren van u als (auto)mobilist.

Wat is uw geslacht?
- ○ Man
- ○ Vrouw

Wat is uw leeftijd? [ ]

Wat is uw hoogst voltooide opleiding?
- ○ Basis onderwijs/Lagere school
- ○ Lager onderwijs (LBO, MAVO, VMBO)
- ○ Middelbaar onderwijs (MBO, HAVO, VWO)
- ○ Hoger onderwijs (HBO, Universiteit)
- ○ Overig
- ○ Onbekend

Wat is de grootte van uw huishouden?
- ○ Een persoon
- ○ Twee personen
- ○ Drie personen of meer

Bent u in het bezit van een rijbewijs?
- ○ Ja
- ○ Nee

Zo ja, hoeveel jaar? [ ]

Heeft u de beschikking over een of meerdere auto's?
- ○ Ja
- ○ Nee

Zo ja, hoeveel? [ ]

Is een van deze auto's een lease auto?
- ○ Ja
- ○ Nee

Hoeveel kilometer rijdt u gemiddeld per jaar?
- ○ 0-10.000 kilometer
- ○ 10.000-20.000 kilometer
- ○ 20.000-30.000 kilometer
- ○ Meer dan 30.000 kilometer

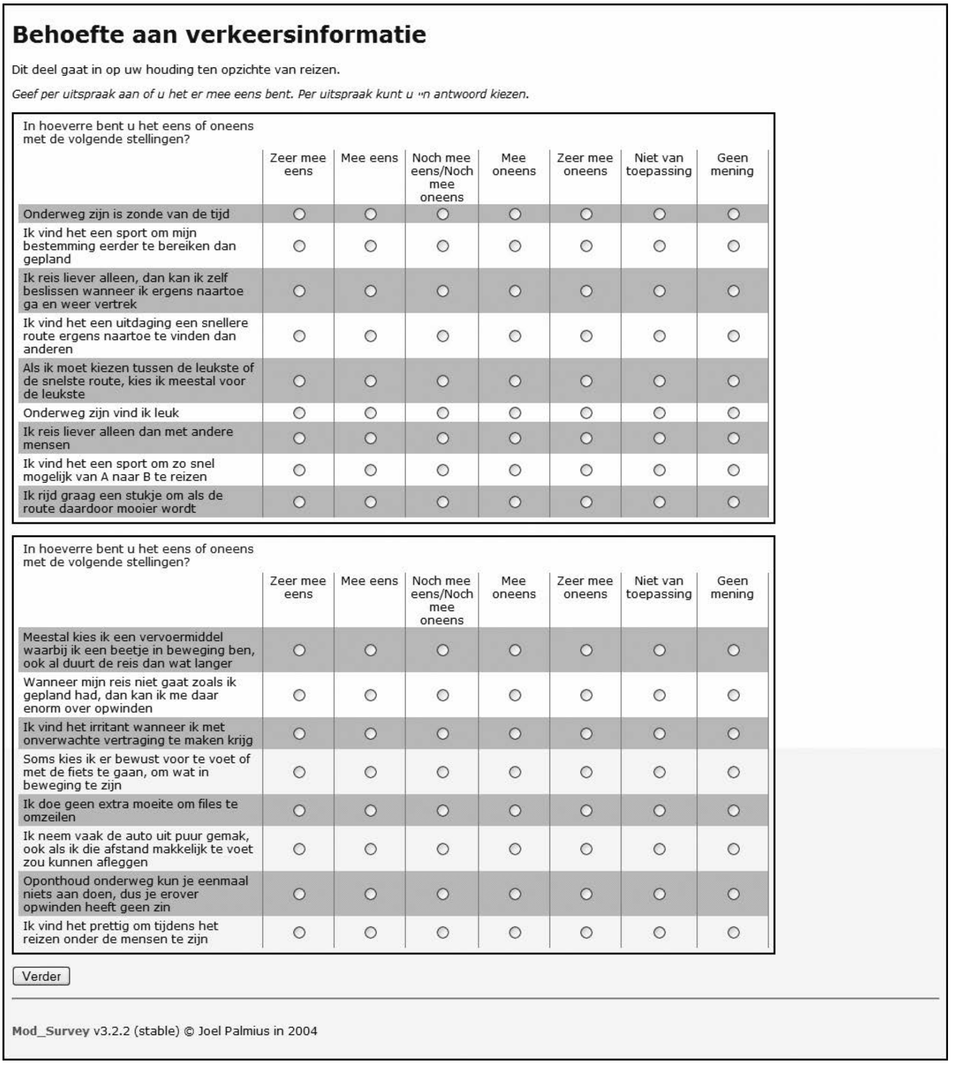
Hoeveel van deze kilometers rijdt u op een ... (totaal 100%)
- [ ] % (auto)snelweg
- [ ] % provinciale weg
- [ ] % stedelijke weg

Waarvoor gebruikt u uw auto het meest?
- ○ Van huis naar werk of vice-versa
- ○ Voor mijn werk (dus geen woon-werk)
- ○ Voor privé reizen

[Verder]

Mod_Survey v3.2.2 (stable) © Joel Palmius in 2004

Figure A.2: *survey1.survey*



Behoefte aan verkeersinformatie

Dit deel gaat in op uw houding ten opzichte van reizen.

Geef per uitspraak aan of u het er mee eens bent. Per uitspraak kunt u ··n antwoord kiezen.

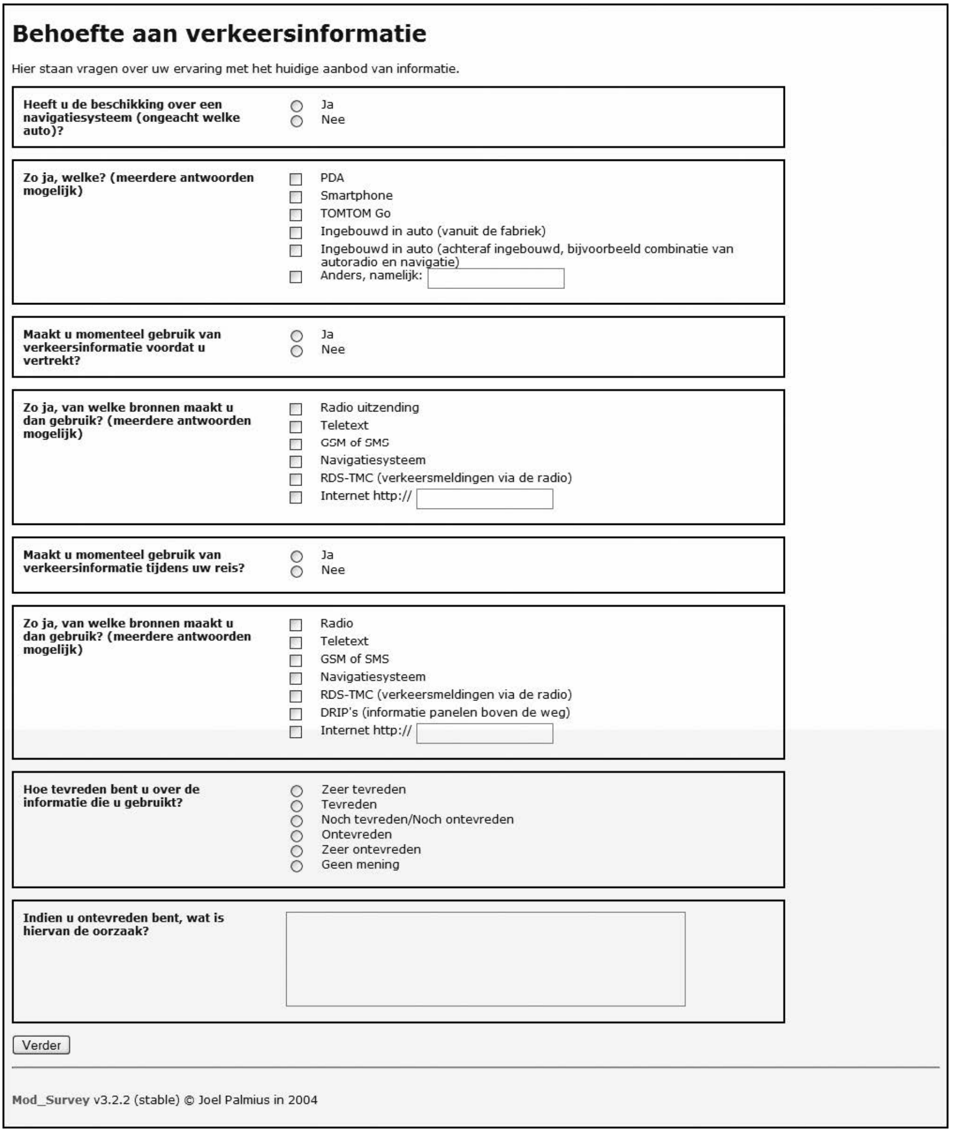
In hoeverre bent u het eens of oneens met de volgende stellingen?	Zeer mee eens	Mee eens	Noch mee eens/Noch mee oneens	Mee oneens	Zeer mee oneens	Niet van toepassing	Geen mening
Onderweg zijn is zonde van de tijd	○	○	○	○	○	○	○
Ik vind het een sport om mijn bestemming eerder te bereiken dan gepland	○	○	○	○	○	○	○
Ik reis liever alleen, dan kan ik zelf beslissen wanneer ik ergens naartoe ga en weer vertrek	○	○	○	○	○	○	○
Ik vind het een uitdaging een snellere route ergens naartoe te vinden dan anderen	○	○	○	○	○	○	○
Als ik moet kiezen tussen de leukste of de snelste route, kies ik meestal voor de leukste	○	○	○	○	○	○	○
Onderweg zijn vind ik leuk	○	○	○	○	○	○	○
Ik reis liever alleen dan met andere mensen	○	○	○	○	○	○	○
Ik vind het een sport om zo snel mogelijk van A naar B te reizen	○	○	○	○	○	○	○
Ik rijd graag een stukje om als de route daardoor mooier wordt	○	○	○	○	○	○	○

In hoeverre bent u het eens of oneens met de volgende stellingen?	Zeer mee eens	Mee eens	Noch mee eens/Noch mee oneens	Mee oneens	Zeer mee oneens	Niet van toepassing	Geen mening
Meestal kies ik een vervoermiddel waarbij ik een beetje in beweging ben, ook al duurt de reis dan wat langer	○	○	○	○	○	○	○
Wanneer mijn reis niet gaat zoals ik gepland had, dan kan ik me daar enorm over opwinden	○	○	○	○	○	○	○
Ik vind het irritant wanneer ik met onverwachte vertraging te maken krijg	○	○	○	○	○	○	○
Soms kies ik er bewust voor te voet of met de fiets te gaan, om wat in beweging te zijn	○	○	○	○	○	○	○
Ik doe geen extra moeite om files te omzeilen	○	○	○	○	○	○	○
Ik neem vaak de auto uit puur gemak, ook als ik die afstand makkelijk te voet zou kunnen afleggen	○	○	○	○	○	○	○
Oponthoud onderweg kun je eenmaal niets aan doen, dus je erover opwinden heeft geen zin	○	○	○	○	○	○	○
Ik vind het prettig om tijdens het reizen onder de mensen te zijn	○	○	○	○	○	○	○

Verder

Mod_Survey v3.2.2 (stable) © Joel Palmius in 2004

Figure A.3: *survey2.survey*



Behoefte aan verkeersinformatie

Hier staan vragen over uw ervaring met het huidige aanbod van informatie.

Heeft u de beschikking over een navigatiesysteem (ongeacht welke auto)?
- Ja
- Nee

Zo ja, welke? (meerdere antwoorden mogelijk)
- PDA
- Smartphone
- TOMTOM Go
- Ingebouwd in auto (vanuit de fabriek)
- Ingebouwd in auto (achteraf ingebouwd, bijvoorbeeld combinatie van autoradio en navigatie)
- Anders, namelijk:

Maakt u momenteel gebruik van verkeersinformatie voordat u vertrekt?
- Ja
- Nee

Zo ja, van welke bronnen maakt u dan gebruik? (meerdere antwoorden mogelijk)
- Radio uitzending
- Teletext
- GSM of SMS
- Navigatiesysteem
- RDS-TMC (verkeersmeldingen via de radio)
- Internet http://

Maakt u momenteel gebruik van verkeersinformatie tijdens uw reis?
- Ja
- Nee

Zo ja, van welke bronnen maakt u dan gebruik? (meerdere antwoorden mogelijk)
- Radio
- Teletext
- GSM of SMS
- Navigatiesysteem
- RDS-TMC (verkeersmeldingen via de radio)
- DRIP's (informatie panelen boven de weg)
- Internet http://

Hoe tevreden bent u over de informatie die u gebruikt?
- Zeer tevreden
- Tevreden
- Noch tevreden/Noch ontevreden
- Ontevreden
- Zeer ontevreden
- Geen mening

Indien u ontevreden bent, wat is hiervan de oorzaak?

Verder

Mod_Survey v3.2.2 (stable) © Joel Palmius in 2004

Figure A.4: *survey3.survey*

Behoefte aan verkeersinformatie

Hieronder wordt u gevraagd aan te geven welk informatie u zou willen ontvangen in 11 verschillende (hypothetische) situaties. Voor elke situatie is een aantal kenmerken gegeven. Het gaat om:

- Situatie die voorkomt op het wegennet;
- Het motief waarmee u op dat moment reist (zakelijk wanneer u voor uw werk reist (niet woon-werk verkeer), privé als u in uw vrije tijd reist, bijvoorbeeld voor familiebezoek);
- De periode waarin u reist;
- Hoe goed u het wegennetwerk "kent" (bekend met alternatieve routes of juist niet);
- Hoe goed u het verkeer "kent" (bekend met de normale, dagelijkse files, of juist niet);
- De normale duur van uw reis naar uw bestemming.

Per situatie kunt u aangeven welke informatie u het liefst zou willen ontvangen (één antwoord mogelijk). Mocht u andere informatie willen ontvangen dan de voorgestelde opties, dan kunt u dat aangeven.

Welke informatie zou u willen ontvangen in deze situatie?

*U bent in de **ochtendspits** op weg naar uw **werk**. Uw reistijd is normaal gesproken **meer dan 30 minuten**. U gebruikt deze route **dagelijks** en **kent de omgeving goed**. U **weet** hoe de **normale verkeerssituatie** er tijdens de ochtendspits uitziet en u kent een aantal **alternatieve routes**. Vlak na vertrek wordt gemeld dat er een **ongeval** is gebeurd op uw dagelijkse route.*

- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Advies voor de mooiste route
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Advies voor de kortste route
- ○ Advies voor de snelste route
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Anders, namelijk: ________

Welke informatie zou u willen ontvangen in deze situatie?

*U bent in de **ochtendspits** op weg naar een **klant van uw bedrijf**. Uw reistijd is normaal gesproken **meer dan 30 minuten**. U **kent de omgeving en de route goed**. U **weet** hoe de **normale verkeerssituatie** er tijdens de ochtendspits uitziet en u kent een aantal **alternatieve routes**. Vlak na vertrek wordt gemeld dat er een **ongeval** is gebeurd op uw route.*

- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Advies voor de kortste route
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Advies voor de snelste route
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Advies voor de mooiste route
- ○ Anders, namelijk: ________

Welke informatie zou u willen ontvangen in deze situatie?

*U bent in de **ochtendspits** op weg naar een **klant van uw bedrijf**. Uw reistijd is normaal gesproken **meer dan 30 minuten**. De **normale verkeerssituatie** tijdens de ochtendspits is voor u **onbekend**. U weet ook **geen alternatieve routes** naar uw bestemming. Vlak na vertrek wordt gemeld dat er een **ongeval** is gebeurd op uw route.*

- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Advies voor de mooiste route
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Advies voor de snelste route
- ○ Advies voor de kortste route
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Anders, namelijk: ________

Welke informatie zou u willen ontvangen in deze situatie?

*U bent in de **ochtendspits** op weg naar bijvoorbeeld **familie of een winkelcentrum**. Uw reistijd is normaal gesproken **minder dan 30 minuten**. De **normale verkeerssituatie** tijdens de ochtendspits is voor u **onbekend**. U weet ook **geen alternatieve routes** naar uw bestemming. Vlak na vertrek wordt gemeld dat er een **ongeval** is gebeurd op uw route.*

- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Advies voor de kortste route
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Advies voor de snelste route
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Advies voor de mooiste route
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Anders, namelijk: ________

(a)

Figure A.5: *survey4.survey part 1*

Welke informatie zou u willen ontvangen in deze situatie?

U bent in de ***ochtendspits*** *op weg naar uw* ***werk****. Uw reistijd is normaal gesproken* ***meer dan 30 minuten****. U gebruikt deze route* ***dagelijks*** *en* ***kent de omgeving goed****. U* ***weet*** *hoe de* ***normale verkeerssituatie*** *er tijdens de ochtendspits uitziet en u kent een aantal* ***alternatieve routes****. Het is de* ***eerste dag*** *dat op uw route* ***grootschalige wegwerkzaamheden*** *plaats vinden.*

- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Advies voor de snelste route
- ○ Advies voor de kortste route
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Advies voor de mooiste route
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Anders, namelijk: [ ]

Welke informatie zou u willen ontvangen in deze situatie?

U bent in de ***ochtendspits*** *op weg naar een* ***klant van uw bedrijf****. Uw reistijd is normaal gesproken* ***meer dan 30 minuten****. U* ***kent de omgeving en de route goed****. U* ***weet*** *hoe de* ***normale verkeerssituatie*** *er tijdens de ochtendspits uitziet en u kent een aantal* ***alternatieve routes****. Het is de* ***eerste dag*** *dat op uw route* ***grootschalige wegwerkzaamheden*** *plaats vinden.*

- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Advies voor de snelste route
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Advies voor de kortste route
- ○ Advies voor de mooiste route
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Anders, namelijk: [ ]

Welke informatie zou u willen ontvangen in deze situatie?

U bent in de ***ochtendspits*** *op weg naar een* ***klant van uw bedrijf****. Uw reistijd is normaal gesproken* ***meer dan 30 minuten****. De* ***normale verkeerssituatie*** *tijdens de ochtendspits is voor u* ***onbekend****. U weet ook* ***geen alternatieve routes*** *naar uw bestemming. Het is de* ***eerste dag*** *dat op uw route* ***grootschalige wegwerkzaamheden*** *plaats vinden.*

- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Advies voor de mooiste route
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Advies voor de kortste route
- ○ Advies voor de snelste route
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Anders, namelijk: [ ]

Welke informatie zou u willen ontvangen in deze situatie?

U bent in de ***ochtendspits*** *op weg naar bijvoorbeeld* ***familie of een winkelcentrum****. Uw reistijd is normaal gesproken* ***meer dan 30 minuten****. De* ***normale verkeerssituatie*** *tijdens de ochtendspits is voor u* ***onbekend****. U weet ook* ***geen alternatieve routes*** *naar uw bestemming. Het is de* ***eerste dag*** *dat op uw route* ***grootschalige wegwerkzaamheden*** *plaats vinden.*

- ○ Advies voor de mooiste route
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Advies voor de snelste route
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Advies voor de kortste route
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Anders, namelijk: [ ]

(b)

Figure A.5: *survey4.survey part 2*

Welke informatie zou u willen ontvangen in deze situatie?

U bent in de ***avondspits*** *op weg naar een* ***groot evenement****, zoals een popconcert. Uw reistijd is normaal gesproken* ***meer dan 30 minuten****. De* ***normale verkeerssituatie*** *tijdens de avondspits is voor u* ***onbekend****. Omdat u wel dagelijks het wegennetwerk in de omgeving gebruikt weet u wel een aantal* ***alternatieve routes*** *naar uw bestemming.*

- ○ Advies voor de kortste route
- ○ Advies voor de snelste route
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Advies voor de mooiste route
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Anders, namelijk: [ ]

Welke informatie zou u willen ontvangen in deze situatie?

U bent in het ***weekend*** *op weg naar een* ***groot evenement****, zoals een festival of Sail Amsterdam. Uw reistijd is normaal gesproken* ***meer dan 30 minuten****. De* ***normale verkeerssituatie*** *tijdens het weekend is voor u* ***onbekend****. Omdat u wel dagelijks het wegennetwerk in de omgeving gebruikt weet u wel een aantal* ***alternatieve routes*** *naar uw bestemming.*

- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Advies voor de kortste route
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Advies voor de snelste route
- ○ Advies voor de mooiste route
- ○ Anders, namelijk: [ ]

Welke informatie zou u willen ontvangen in deze situatie?

U bent in de ***avondspits*** *op weg naar uw* ***huis*** *vanaf uw werk. Uw reistijd is normaal gesproken* ***meer dan 30 minuten****. U gebruikt deze route* ***dagelijks*** *en* ***kent de omgeving goed****. U weet hoe de normale verkeerssituatie er tijdens de avondspits uitziet en u kent een aantal* ***alternatieve routes****. Het is de* ***eerste dag*** *dat op uw route* ***grootschalige wegwerkzaamheden*** *plaats vinden.*

- ○ Advies voor de mooiste route
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (advies route)
- ○ Advies voor parkeerplaats met route hier naartoe
- ○ Filelokatie, filelengte, file oorzaak en verwachte duur van de file (gehele netwerk)
- ○ Advies voor de kortste route
- ○ Alternatieve vervoermiddelen naar de bestemming
- ○ Filelokatie en filelengte (gehele netwerk)
- ○ Bezienswaardigheden in de omgeving met routes hier naartoe
- ○ Advies voor de snelste route
- ○ Verwachte aankomsttijd en 95% zekerheidsmarges (vaste route)
- ○ Tankstations/hotels/restaurants in de omgeving met routes hier naartoe
- ○ Anders, namelijk: [ ]

[Verder]

(c)

Figure A.5: *survey4.survey part 3*

Behoefte aan verkeersinformatie

Op deze pagina wordt u gevraagd om voor dezelfde situatie 9 keer een keus te maken uit 4 soorten informatie of te kiezen voor geen informatie. De 4 soorten informatie blijven telkens hetzelfde, maar variëren telkens wel qua kenmerken, zoals kosten en betrouwbaarheid. Lees de beschrijvingen goed door.

De 4 soorten informatie zijn:

- Informerend: informatie over de lengte en lokaties van files evenals mogelijke oorzaken; u moet zelf nog een keuze maken voor de "beste" route
- Adviserend: "algemeen" advies voor de snelste route
- Persoonlijk informerend en adviserend: persoonlijk advies voor een beste route, gebaseerd op de kenmerken die voor u van belang zijn (snel, kortste afstand, veel/weinig snelweg, etc.) met daarbij informatie over points of interest, zoals parkeerplaatsen en restaurants, die voor u interessant zijn
- Voorspellend: een persoonlijk advies, maar nu ook gebaseerd op een voorspelling van toekomstige verkeerssituaties

De kenmerken waarin de soorten informatie variëren zijn:

- Kosten: de totale kosten van het gebruik, inclusief aanschaf, software en dataverbruik
- Tijdigheid: de periode tussen het meten van de verkeerssituatie en het doorgeven van de informatie aan de reiziger, veroorzaakt door het interpreteren en versturen van de informatie
- Betrouwbaarheid: de mate waarin de verstuurde informatie of het verstuurde advies klopt met de werkelijke situatie

De situatie waar het in **alle** gevallen om gaat, is de volgende:
U bent op weg naar uw werk, in de ochtendspits. U gaat dagelijks naar uw werk en dat al een paar jaar lang. U bent bekend met de files die er kunnen staan en weet een aantal mogelijkheden om uw bestemming te bereiken. De reis van uw huis naar uw werk duurt altijd langer dan een half uur, maar de reistijd is geen enkele dag gelijk. De informatie die u opvraagt krijgt u zodra u uw reis bent begonnen. Het is dus niet meer mogelijk om het vertrektijdstip aan te passen. Er wordt van u verwacht dat u op een vastgesteld tijdstip aanwezig bent op uw werk.

U kunt nu voor deze situatie 9 keer kiezen uit de vier soorten informatie, of er voor kiezen geen informatie te ontvangen.

Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €0,55 per reis • ongeveer 5 minuten vertraging • klopt wel eens	• €0,55 per reis • geen vertraging • klopt wel eens	• €0,10 per reis • ongeveer 15 minuten vertraging • klopt altijd	• €0,55 per reis • geen vertraging • klopt altijd	
○	○	○	○	○

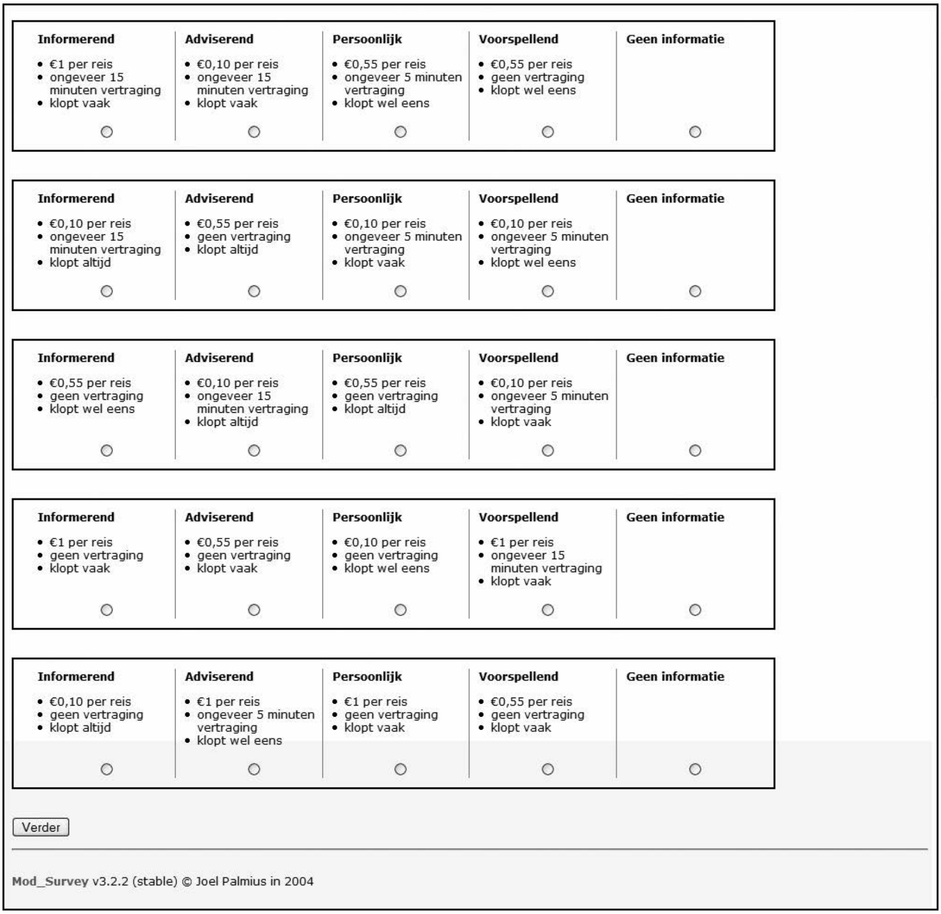
Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €1 per reis • ongeveer 5 minuten vertraging • klopt vaak	• €1 per reis • ongeveer 5 minuten vertraging • klopt altijd	• €1 per reis • ongeveer 15 minuten vertraging • klopt wel eens	• €0,10 per reis • ongeveer 5 minuten vertraging • klopt altijd	
○	○	○	○	○

Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €0,10 per reis • ongeveer 5 minuten vertraging • klopt altijd	• €0,10 per reis • ongeveer 15 minuten vertraging • klopt vaak	• €0,55 per reis • ongeveer 15 minuten vertraging • klopt vaak	• €1 per reis • ongeveer 15 minuten vertraging • klopt altijd	
○	○	○	○	○

Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €0,55 per reis • ongeveer 15 minuten vertraging • klopt wel eens	• €1 per reis • ongeveer 5 minuten vertraging • klopt vaak	• €1 per reis • ongeveer 5 minuten vertraging • klopt altijd	• €1 per reis • ongeveer 15 minuten vertraging • klopt wel eens	
○	○	○	○	○

(a)

Figure A.6: *DC1.survey part 1*



Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €1 per reis • ongeveer 15 minuten vertraging • klopt vaak	• €0,10 per reis • ongeveer 15 minuten vertraging • klopt vaak	• €0,55 per reis • ongeveer 5 minuten vertraging • klopt wel eens	• €0,55 per reis • geen vertraging • klopt wel eens	

Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €0,10 per reis • ongeveer 15 minuten vertraging • klopt altijd	• €0,55 per reis • geen vertraging • klopt altijd	• €0,10 per reis • ongeveer 5 minuten vertraging • klopt vaak	• €0,10 per reis • ongeveer 5 minuten vertraging • klopt wel eens	

Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €0,55 per reis • geen vertraging • klopt wel eens	• €0,10 per reis • ongeveer 15 minuten vertraging • klopt altijd	• €0,55 per reis • geen vertraging • klopt altijd	• €0,10 per reis • ongeveer 5 minuten vertraging • klopt vaak	

Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €1 per reis • geen vertraging • klopt vaak	• €0,55 per reis • geen vertraging • klopt vaak	• €0,10 per reis • geen vertraging • klopt wel eens	• €1 per reis • ongeveer 15 minuten vertraging • klopt vaak	

Informerend	Adviserend	Persoonlijk	Voorspellend	Geen informatie
• €0,10 per reis • geen vertraging • klopt altijd	• €1 per reis • ongeveer 5 minuten vertraging • klopt wel eens	• €1 per reis • geen vertraging • klopt vaak	• €0,55 per reis • geen vertraging • klopt vaak	

Verder

Mod_Survey v3.2.2 (stable) © Joel Palmius in 2004

(b)

Figure A.6: *DC1.survey part 2*

Behoefte aan verkeersinformatie

Mede op basis van de resultaten van deze enquête wordt een onderzoek gedaan naar de keuze van bepaalde routes. Dit vervolgonderzoek wordt uitgevoerd op de Universiteit Twente in Enschede.

Zou u mee willen werken aan dit vervolgonderzoek? Ja / Nee

Wilt u meedoen met de verloting van de cadeaubon? Ja / Nee

Indien u op één of allebei van bovenstaande vragen met ja heeft geantwoord, geeft u dan onderstaande gegevens. Er wordt dan te zijnertijd contact met u opgenomen.

Voorletters

Achternaam

Tussenvoegsel

Straat

Nummer

Postcode

Plaats

Telefoonnummer

Email adres

Verder

Mod_Survey v3.2.2 (stable) © Joel Palmius in 2004

Figure A.7: *survey6.survey*

Behoefte aan verkeersinformatie

Bedankt voor uw medewerking aan deze enquête. Mocht u nog opmerkingen hebben over de enquête dan kunt u die hieronder kwijt. Vragen zijn ook altijd welkom.

Opmerkingen/Vragen

Einde

Mod_Survey v3.2.2 (stable) © Joel Palmius in 2004

Figure A.8: *survey7.survey*

Appendix B

Current information usage

Table B.1: *Current information usage for different genders*

Question	Answer	Male	Female
Do you possess a navigation system?	yes	28.5%	21.7%
	no	71.5%	78.3%
Do you usage traffic information before leaving?	yes	63.7%	52.7%
	no	36.3%	47.3%
Do you usage traffic information during your trip?	yes	79.6%	66.4%
	no	20.4%	33.6%

Table B.2: *Current information usage for different agegroups*

Question	Answer	18-24	25-44	45-64	65>
Do you possess a navigation system?	yes	12.5%	25.1%	30.5%	17.9%
	no	87.5%	74.9%	69.5%	82.1%
Do you usage traffic information before leaving?	yes	50.0%	56.4%	62.8%	82.1%
	no	50.0%	43.6%	37.2%	17.9%
Do you usage traffic information during your trip?	yes	55.0%	74.9%	77.1%	78.6%
	no	45.0%	25.1%	22.9%	21.4%

Table B.3: *Current information usage for differen education levels*

Question	Answer	primary	lower secondary	higher secondary	higher and university
Do you possess a navigation system?	yes	33.3%	37.0%	24.4%	25.2%
	no	66.7%	63.0%	75.6%	74.8%
Do you usage traffic information before leaving?	yes	66.7%	43.5%	59.2%	61.3%
	no	33.3%	56.5%	40.8%	38.7%
Do you usage traffic information during your trip?	yes	50.0%	60.9%	68.2%	80.6%
	no	50.0%	39.1%	31.8%	19.4%

Table B.4: *Current information usage for different kilometers per year*

Question	Answer	0-10.000 km	10-20.000 km	20-30.000 km	>30.000 km
Do you possess a navigation system?	yes	33.3%	37.0%	24.4%	25.2%
	no	66.7%	63.0%	75.6%	74.8%
Do you usage traffic information before leaving?	yes	66.7%	43.5%	59.2%	61.3%
	no	33.3%	56.5%	40.8%	38.7%
Do you usage traffic information during your trip?	yes	50.0%	60.9%	68.2%	80.6%
	no	50.0%	39.1%	31.8%	19.4%

Table B.5: *Current information usage for different motives for car use*

Question	Answer	Commute	Business	Recreation
Do you possess a navigation system?	yes	26.7%	41.7%	20.0%
	no	73.3%	58.3%	80.0%
Do you usage traffic information before leaving?	yes	56.7%	66.7%	59.3%
	no	43.3%	33.3%	40.7%
Do you usage traffic information during your trip?	yes	73.3%	87.5%	71.2%
	no	26.7%	12.5%	28.8%

Table B.6: *Current information usage for different attitudes towards mobility*

Question	Answer	Accepter	Deliberate	Competitor	Conscious	Enjoyers
Do you possess a navigation system?	yes	21.0%	24.2%	32.9%	25.7%	32.0%
	no	79.0%	75.8%	67.1%	74.3%	68.0%
Do you usage traffic information before leaving?	yes	52.1%	66.2%	65.9%	71.6%	48.0%
	no	47.9%	33.8%	34.1%	28.4%	52.0%
Do you usage traffic information during your trip?	yes	69.5%	79.0%	83.5%	73.0%	68.0%
	no	30.5%	21.0%	16.5%	27.0%	32.0%

Appendix C

AIDA RCS Client screens

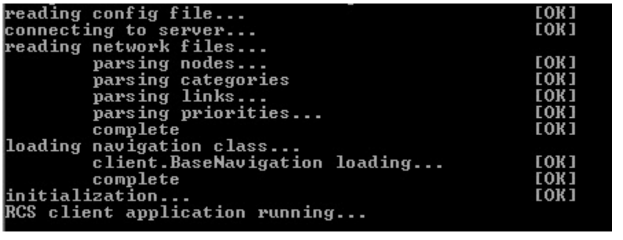


Figure C.1: *AIDA RCS Client startup sequence*

Figure C.2: *AIDA RCS Client login screen*

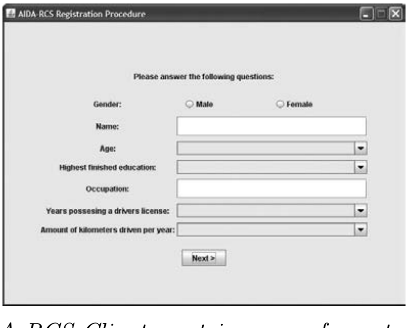


Figure C.3: *AIDA RCS Client pre trip screen for entering personal data 1*

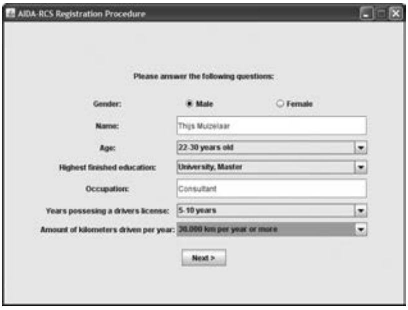


Figure C.4: *AIDA RCS Client pre trip screen for entering personal data 2*

Figure C.5: *AIDA RCS Client scenario selection screen 1*

Figure C.6: *AIDA RCS Client scenario selection screen 2*

Figure C.7: *AIDA RCS Client guidance selection screen*

Figure C.8: *AIDA RCS Client start screen*

Figure C.9: *AIDA RCS Client start dashboard screen*

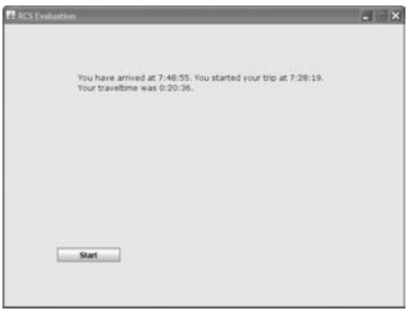


Figure C.10: *AIDA RCS Client arrival screen*

Figure C.11: *AIDA RCS Client evaluation screen 1*

Figure C.12: *AIDA RCS Client evaluation screen 2*

Figure C.13: *AIDA RCS Client evaluation screen 3*

Figure C.14: *AIDA RCS Client finish screen*

Appendix D

Participant introduction

Beste deelnemer,

Welkom bij het routekeuze experiment. Alvast hartelijk bedankt voor je deelname. Door je deelname maak je het mogelijk om de verkeersinformatie in de toekomst beter toe te spitsen op bepaalde verkeerssituaties. De verkeersinformatie kan ook persoonlijker worden gemaakt, uitgaande van je eigen voorkeuren.

In dit experiment ga je een aantal keer dezelfde reis maken. Tijdens die reis moet je bij elke kruising een nieuwe keuze maken. De verkeerssituaties die je onderweg tegenkomt zijn niet telkens dezelfde. Daarom is het van belang goed op te letten en bewust je keuze te maken.

Voordat je met het experiment kunt beginnen moet je inloggen op het systeem. Onderstaande screenshot laat dit eerste scherm zien. Heb je nog geen reis gemaakt, kies dan voor geen geregistreerde gebruiker. Heb je al een keer een reis gemaakt, dan heb je je naam en registration ID gekregen van het systeem.

Voor de registratie moet je een aantal persoonlijke kenmerken invullen. In onderstaand scherm zie je welke dat precies zijn. Ben je klaar, druk dan op submit.

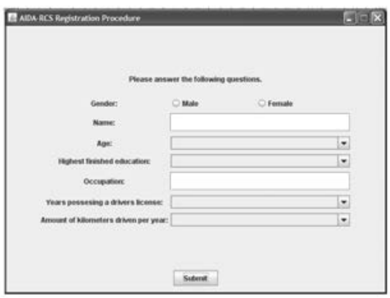


Figure D.1: *Participant introduction page 1*

Nadat je bent geregistreerd moet je kiezen voor het type informatie dat je ontvangt. Op je scenarioblad staat welke informatie je moet kiezen tijdens de reis die je dan gaat maken. Je kunt hier kiezen uit:

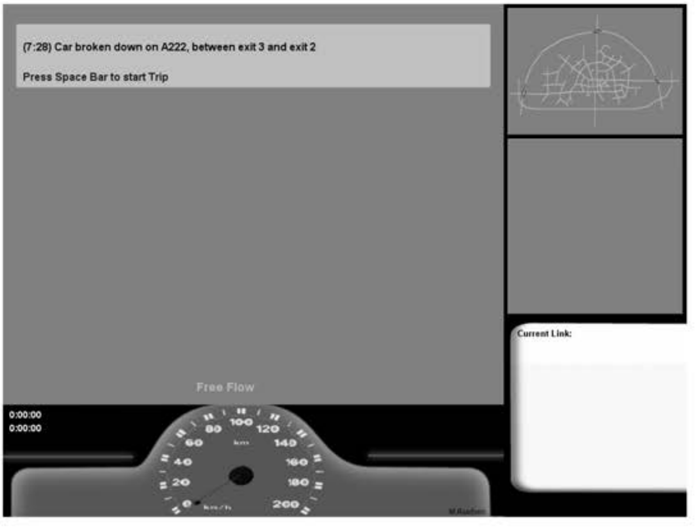
- Geen informatie
- Verkeers en speciale gebeurtenis informatie
- Snelste route advies en verkeers en speciale gebeurtenis informatie.

Na de keuze voor de verkeersinformatie moet je je scenario kiezen. In dit scenario staat met welk motief je gaat reizen, waar je vandaan vertrekt en waar je naar toe gaat. Ook staat er hoe laat je normaal vertrekt en moet aankomen op je bestemming.

De volgende stap geeft nog een kort overzicht van de situatie die je kunt tegenkomen. Via Submit kom je in de experiment omgeving.

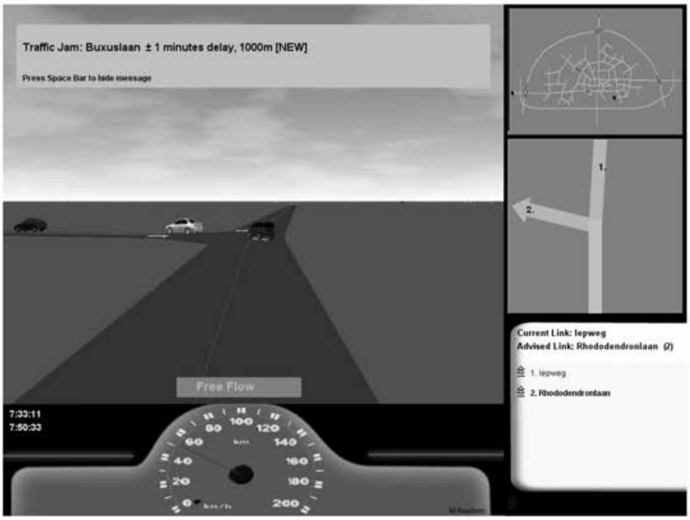
Figure D.2: *Participant introduction page 2*

Bij het starten van elke rit ziet de experiment omgeving eruit zoals hieronder.



Linksboven zie je hier een verkeersbericht. Bij het starten van elke rit wordt de op dat moment aanwezige verkeersinformatie getoond (uiteraard alleen als daar eerder voor is gekozen). Dat kan gaan om (zoals in het voorbeeld) informatie over een ongeval, of over files. Met de spatiebalk kun je de rit dan daadwerkelijk starten.

Figure D.3: *Participant introduction page 3*



Is de rit eenmaal gestart, dan verandert het scherm, zoals hierboven is te zien. Linksboven wordt weer de verkeersinformatie weergegeven (nu een nieuw ontstane file). Deze is met behulp van de spatiebalk weg te halen. Ook is nu een beeld te zien van de eerstvolgende kruising, of indien er een grote afstand is tussen twee kruisingen, een beeld van halverwege de weg. Onderaan in dat beeld is te zien hoe druk het nu is op de weg waar je bent. Dat kan zijn:

- Free Flow (groen): vrije doorstroming
- Medium Traffic (oranje): het is al wat drukker
- Heavy Traffic (rood): het is erg druk.

Daaronder zie je je dashboard, met daarin de huidige snelheid die je rijdt. Links op je dashboard staat de huidige tijd, en indien er gebruik wordt gemaakt van een snelste route advies, dan staat daaronder de verwachte aankomsttijd.

De rechterkant van het scherm biedt je meer informatie. Het bovenste vlakje is een kaart van de stad waarin je je bevind. De blauwe stipjes zijn je herkomst en bestemming. Je eigen lokatie wordt aangegeven met een wit stipje. De kleuren van elk wegvak komen weer overeen met de drukte. Hier kun je dus files min of meer terugvinden. De kaart draait echter niet mee met de richting die je oprijdt.

Figure D.4: *Participant introduction page 4*

Het volgende vlakje geeft de eerstvolgende kruising weer. In bovenstaand voorbeeld heb je twee opties, die zijn genummerd. De drukte van de twee opties wordt weer door middel van een kleur aangegeven. Bovenstaand voorbeeld geeft ook een advies aan, in de vorm van een pijl. Als er geen gebruik wordt gemaakt van een snelste route advies, dan worden alle opties zonder pijl getekend. Let op dat deze weergave wel meedraait met de rijrichting, in tegenstelling tot de kaart van het wegennet.

Het derde vlak geeft eerst aan wat je huidige weg is. Als er weer gebruikt wordt gemaakt van een snelste route advies, dan staat hier ook de straatnaam van de weg die wordt geadviseerd, anders blijft dat leeg. Daaronder staan de opties waar je uit kunt kiezen, nu met een straatnaam erbij. De kleine gekleurde driehoekjes geven wederom de drukte weer. Als je een keuze hebt gemaakt (met behulp van een cijfer op het toetsenbord) dan wordt de gekozen optie blauw gekleurd. Je kunt je keuze altijd nog wijzigen, totdat de blauwe balk onderaan is verdwenen.

Het laatste onderdeel is deze blauwe balk. Dit geeft de tijd weer die je nog hebt om je keuze te maken. De balk begint helemaal vol en loopt terug naar links. De periode die je hebt om te kiezen hangt af van de afstand tot de volgende kruising en de snelheid die je kunt rijden.

Hiermee eindigt de introductie van de omgeving. Neem hierna de kaart goed door en bekijk je scenario blad. Mocht je nog vragen hebben, stel ze dan gerust.

Figure D.5: *Participant introduction page 5*

Appendix E

Experiment log

Logblad

Naam	
Registration ID	
Scenario	2

Welkomstblad gegeven	
Straat/snelheid kaarten gegeven	
Scenario kaart gegeven	

Rit	SimID	TripID	Scenario	Seed	Accidents	Info	Restarts
0			1	2	Geen	3	
1			2	2	Geen	1	
2			2	4	Geen	1	
3			2	2	1	1	
4			2	4	2	1	
5			2	2	1	2	
6			2	2	2	2	
7			2	4	1	3	
8			2	4	2	3	

Figure E.1: *Experiment log*

Appendix F

Segmented route choice models

Table F.1: *Segmented base model for the different participants for both datasets*

Variable	Complete	Strategic
ASC1 (right)	0	0
ASC2 (straight)	0.573	−0.010[a]
ASC3 (left)	−0.165[a]	0.041[a]
$\beta_{FFTT(\lambda)}$	−2.07	−5.30[a]
$\beta_{\log D}$	−0.623	−0.604
β_{JC}	−0.215	−0.183
β_C	0.513	
β_I	−0.742[a]	−1.32[a]
β_{TM}	−4.49	
β_A	1.43	1.55
λ	0.385	0.054[a]
μ_1	1	1
μ_2	1.65[a]	0.584[a]
μ_3	2.28[a]	1.40[a]
μ_4	0.925[a]	0.590[a]
μ_5	0.538	0.501
μ_6	1.19[a]	1.33[a]
μ_7	2.07	1.20[a]
μ_8	0.767[a]	0.508
μ_9	1.27[a]	0.875[a]
μ_{10}	1.77[a]	0.952[a]
μ_{11}	0.884[a]	0.757[a]
μ_{12}	1.33[a]	1.05[a]
μ_{13}	1.27[a]	1.04[a]
μ_{14}	0.784[a]	0.509
μ_{15}	0.357	0.416
μ_{16}	0.462	0.504
μ_{17}	0.467	0.309
μ_{18}	0.608	0.367
μ_{19}	1.08[a]	0.882[a]
μ_{20}	0.801[a]	0.646[a]
μ_{21}	1.47[a]	0.925[a]
μ_{22}	0.905[a]	0.772[a]
μ_{23}	0.992[a]	1.18[a]
μ_{24}	1.18[a]	1.27[a]
μ_{25}	1.31[a]	1.11[a]
μ_{26}	0.898[a]	0.751[a]
μ_{27}	1.32[a]	0.872[a]
μ_{28}	1.10[a]	0.984[a]
μ_{29}	0.963[a]	0.786[a]
μ_{30}	1.00[a]	0.665[a]
Number of observations	2517	905
Null log likelihood	−2180.932	−800.432
Final log likelihood	−818.168	−436.607
Likelihood ratio test	2725.528	727.650
Pseudo ρ^2	0.625	0.455
Adjusted pseudo ρ^2	0.607	0.408
Number of parameters	39	37

[a] Non-significant estimate at a 95% level.

Table F.2: *Segmented base model for the complete dataset*

Variable	Scenario	Incident	Information	Trip	Gender	Age	License
ASC1 (right)	0	0	0	0	0	0	0
ASC2 (straight)	0.704	0.524	0.496	0.361	0.554	0.735	0.724
ASC3 (left)	−0.160[a]	−0.136[a]	−0.148[a]	−0.111[a]	−0.189[a]	−0.235[a]	−0.225[a]
$\beta_{FFTT(\lambda)}$	−1.78	−1.11	−1.09	−0.893	−0.992	−1.81	−1.48
$\beta_{\log D}$	−0.562	−0.442	−0.500	−0.367	−0.553	−0.749	−0.744
β_{JC}	−0.223	−0.179	−0.194	−0.144	−0.230	−0.306	−0.322
β_C	0.597	0.426	0.446	0.305	0.572	0.702	0.785
β_I	−1.01	−0.446[a]	−0.658	−0.483	0.593[a]	−0.930	0.796[a]
β_{TM}	−4.65	−3.52	−3.12	−2.51	−3.82	−5.69	−5.53
β_A	1.68	1.17	0.786	0.528	1.32	1.79	1.80
λ	0.513	0.550	0.579	0.537	0.641	0.537	0.582
μ_a	0.617	1.84	1.47	2.01	1.15[a]	0.581	0.663
μ_b	1.01[a]	1.31[a]	1.47[a]	1.47			
μ_c		0.798[a]		1.33[a]			
μ_d		1.21[a]		2.26			
μ_e		1.49		1.83			
μ_f		1.25[a]		1.79			
μ_g				2.53			
Number of observations	2517	2517	2517	2517	2517	2517	2517
Null log likelihood	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932	−2180.932
Final log likelihood	−859.995	−859.508	−867.583	−851.972	−877.159	−850.159	−863.122
Likelihood ratio test	2641.874	2642.849	2626.698	2657.920	2607.546	2661.546	2635.620
Pseudo ρ^2	0.606	0.606	0.602	0.609	0.598	0.610	0.604
Adjusted pseudo ρ^2	0.600	0.599	0.597	0.602	0.593	0.605	0.599
Number of parameters	12	16	12	17	11	11	11

[a] Non-significant estimate at a 95% level.

Table F.3: *Segmented base model for the strategic dataset*

Variable	Scenario	Incident	Information	Trip	Gender	Age	License
ASC1 (right)	0	0	0	0	0	0	0
ASC2 (straight)	-0.017[a]	-0.030[a]	0.037[a]	0.040[a]	-0.036[a]	-0.039[a]	-0.036[a]
ASC3 (left)	-0.029[a]	-0.068[a]	-0.010[a]	-0.013[a]	-0.044[a]	-0.012[a]	-0.025[a]
$\beta_{FFTT(\lambda)}$	-3.68	-3.12	-3.08	-2.85	-3.10	-4.00	-4.01
$\beta_{\log D}$	-0.408	-0.431	-0.405	-0.385	-0.481	-0.627	-0.608
β_{JC}	-0.143	-0.161	-0.101	-0.103	-0.144	-0.193	-0.185
β_I	-1.19	-0.735	-0.684	-0.702	-0.749	-1.02	-0.905
β_A	1.22	0.903	0.580	0.534	1.22	1.44	1.51
λ	0.153[a]	0.045[a]	0.031[a]	0.042[a]	0.144[a]	0.183[a]	0.127[a]
μ_a	0.706	1.19[a]	1.94	1.32[a]	1.12[a]	0.602	0.714
μ_b	1.15[a]	1.40[a]	1.71[a]	0.930[a]			
μ_c		0.762[a]		0.995[a]			
μ_d		1.19[a]		2.02[a]			
μ_e		2.05		2.04			
μ_f		1.08[a]		1.94[a]			
μ_g				1.72[a]			
Number of observations	905	905	905	905	905	905	905
Null log likelihood	-800.432	-800.432	-800.432	-800.432	-800.432	-800.432	-800.432
Final log likelihood	-452.797	-449.820	-448.400	-447.202	-457.781	-449.017	-453.997
Likelihood ratio test	695.270	701.223	704.064	706.460	685.302	702.829	692.870
Pseudo ρ^2	0.434	0.438	0.440	0.441	0.428	0.439	0.433
Adjusted pseudo ρ^2	0.422	0.421	0.427	0.423	0.417	0.428	0.432
Number of parameters	10	14	10	15	9	9	9

[a] Non-significant estimate at a 95% level.

Summary

Land travel has changed dramatically over the course of centuries. The car has become the favorite way of commuting. Increasingly, the road network is incapable of providing sufficient supply (capacity) for the demand (amount of cars) in rush hours resulting in congestion and delay, especially in dense urban areas such as the "Randstad". Congestion leads to significant societal and economic costs, directly because of the lost time and indirectly because congestion increases the unreliability of transport. For example, because of unreliable travel times, a driver plans to leave early. Without congestion, the result of demand and supply results in predictable traffic flows (in terms of travel time, costs, etc.) on the road network. A trip is said to be unreliable if the traffic flows exhibit conditions which were not expected by the traveller, which will often result in a negative experience with the traffic system for the traveller. Reliability thus depends on the expectations of the individual traveller and can be separated into variation and predictability. Large variation and low predictability lead to an unreliable situation. Drivers can deal with unreliability by changing their behaviour and choices, for example by choosing a different route or adapt their departure time.

The reliability of a traffic system can be influenced by Advanced Traveller Information Systems (ATIS). ATIS aim at informing travellers about the current and future state of the transport system, such that the traveller can adjust the expectations or improve their experience with the traffic system. ATIS have been widely studied related to route choice, departure time choice and even mode choice or destination choice. However, most of this research is dedicated to information that targets at an optimal use of the network. Personalized traffic information takes into account the personal preferences, but also the context of the trip at hand. Especially this context can be of importance, as for a trip which is regularly made by a certain traveller the need for information is likely to be different than for a unique trip.

In this thesis, the combination of non-recurrent traffic situations (as a combination of trip context and traffic flows in a road network) and traffic information is studied. This results in the following main objective: ***To gain more insight into the impact of traffic information on route choice behaviour in non-recurrent traffic situations.***

Traffic situations are often depicted from the viewpoint of the road network or its operator. Congestion, as a result of a demand that exceeds supply, then can be labelled as recurrent or non-recurrent. Events such as bad weather can

cause variations in the occurrence of congestion, which makes such congestion non-recurrent. This viewpoint however, ignores the travellers experiencing the congestion. This experience is expected to play a significant role in labelling a traffic situation as recurrent or non-recurrent. In this thesis a framework is developed that integrates both points of view combining congestion and its effects with the relevant elements of the individual traveller. A traveller has a certain degree of familiarity and experience with a trip at hand, in terms of the road network as a mental map and the traffic situations that can occur ranging from very local to high level. Combined with the perceptions of the road network and traffic situations that can occur, this leads to an expectation of the reliability and certainty for the trip to be undertaken. Based on the traffic situations it is possible to determine key elements objectively, such as travel time variability, reliability and predictability. Examples of objective measures range from statistical concepts such as standard deviations to measures which define the extra time necessary to arrive on time in a specified number of occasions. The framework leads to the selection of three different traffic situations for further research: (a) accidents, (b) road works, and (c) large events.

These three non-recurrent traffic situations are the basis for a survey on the needs and preferences of road users for traffic information in these situations. The survey focused on the relationship between situations and personal characteristics of the traveller on one side and the attributes, types and contents of traffic information on the other side. This relation enables the personalization of traffic information depending on the personal characteristics and trip context. A total of 598 Dutch car drivers completed the survey.

The analysis of the survey results indicate that drivers preferred advice on the fastest route, followed by extended traffic information (location, length, cause and expected duration of traffic jams on the complete network) and expected time of arrival, irrespective of the situation. Depending on the situation the preference for traffic information changes and other contents are chosen more often, such as information on a scenic route. The actual traffic situation (e.g. an accident), the motive of the driver (e.g. commute) and the familiarity of the driver are the elements that are of significant influence on the preference of a driver for a specific content for traffic information. Personal characteristics also appeared to be important. The attitude towards mobility, age and current use of traffic information during travel, significantly influence the preferences, irrespective of the situation.

Drivers appeared to prefer advice or personal advice when they were willing to pay for traffic information, out of five different alternatives. One of these alternatives concerned “no information” which was freely available. The other alternatives had to be paid for. In total 60% of all respondents were willing to pay for traffic information and the remaining 40% choose not to acquire any traffic information. In order to determine the (relative) importance of the type and attributes of traffic information, several discrete choice models were estimated. Obviously, costs are the most important attribute of traffic information, with increasing cost reducing the utility of a specific type of traffic information. Second most important is reliability and least important is timeliness. Attitude towards mobility as a determining factor for different groups changed the relative importance of the attributes, but did not influence the order hereof. All these results indicate a need for personalization of traffic information, for contents, types and attributes.

In case of accidents, advice for the fastest route and the extended traffic information (containing information on the location, length, cause and expected duration of congestion) were the most preferred contents of traffic information. Such traffic information is targeted at improving the current trip of a driver in terms of travel time, especially by changing the route choice behaviour. A new route choice simulator was designed and built (the AIDA RCS which stands for Application of Integrated Driver Assistance Route Choice Simulator), based on the ability to incorporate accidents and their effects, together with traffic information and route guidance. The AIDA RCS consists of a micro simulation model for the traffic simulation, a server dealing with data storage and communication with the traffic model, and a client showing the upcoming junction to the participant, together with a dashboard and traffic information and/or route guidance. The participant could indicate the route choice by selecting one of the exits at the upcoming junction. The server stored all route choices made by participants, together with the provided information and data on the traffic situation.

The AIDA RCS was used in an experiment to determine the impact of traffic information and route guidance on route choice behaviour during non-recurrent traffic situations. A combination of an urbanized road network, a morning commute and accidents were chosen as non-recurrent traffic situation. Three combinations of origin and destination were used as a scenario each with specific accidents. 30 participants took part in the experiment. Each participant made 8 trips in a single scenario. The face validity of the AIDA RCS was investigated based on responses of the participants. In general, they reported a realistic experience in the AIDA RCS, with realistic trips. They indicated they would make a similar route choice in real life. This high level of face validity is similar to the validity of other route choice simulators as found in other studies.

The results showed that participants changed routes when first receiving traffic information and are further "improving" their route choice as they receive dynamic route guidance. The expected effect of concentration of participants using the same routes was observed, as less variation between participants occurred. In each scenario one accident was located on what turned out to be the preferred route, as participants rerouted during their trip as they encountered the accident or its consequences. The accidents caused an increase of the travel time by almost 10% with an increase of delay of around 30% when no information was available. Traffic information and dynamic route guidance decreased the travel time by 13% and 6% respectively. Both also reduced the variation of travel time and delay, and the length of the routes used. The individual choices showed that in most cases participants selected the fastest exit at a junction (88%), which increased to 94% in case of dynamic route guidance. Accidents had a significant effect on the amount of choices for the fastest exit, especially in case they were on the preferred route.

The individual choices at the junctions formed the input for an exit choice model. This exit choice model assumes that a driver makes a choice for a route at each junction, independent of the previous choices, which reduces the amount of alternatives to two or three. To assure the exits lead to viable routes, the fastest route starting at the exit was calculated to the destination. This information was used to describe the alternatives a participant choose between. The decision rules for this discrete choice model are based on the principle of utility maximization. This means that individuals choose the alternative with the highest utility in their situation. The multinomial logit model is based on this principle together with specific assumptions about the statistical characteristics unknown elements for the utility of an alternative. The attributes of the alternatives that are used in the model are the freeflow travel time, delay, number of intersections and a dummy variable indicating the exit brings the participant closer to the destination. Of these attributes, freeflow travel time was found to be the most important, but an additional unit of travel time at larger values is of less influence than at small values. This confirms expected behaviour in route choice. Delay was found to be the second most important attribute.

Additional attributes for the traffic situation and traffic information resulted in a significant improvement of the model estimation. The additional attribute of having advice on a certain exit is more important in the choice for an exit than freeflow travel time of this route. The variable for accidents on the route does not have a significant impact on the utility of an alternative, but still is an important indication that a route with accidents is less preferred. Applying models that account for violations of the underlying assumptions of the multinomial logit model, does not change the behavioural implications already found, and do not show a better estimation. A significant difference is found between the scenarios, which indicates that not all combinations of origin and destination are expected to have the same importance for all attributes. These models also showed that the current choice is independent of the choices at previous junctions, and even from previous trips. They also indicate that there is no significant learning effect during the 8 trips of a single participant. An application of the estimated model to a part of the dataset with individual choice which is not used for estimation showed that the predictions are close to the actual choices.

In conclusion, this thesis provides insight in the impact of traffic information on route choice behaviour in non-recurrent traffic situations. The framework for non-recurrent traffic situations provides a broader view which can be used to determine the need and preference for traffic information. The need and preference are shown to differ for different traffic situations and personal characteristics, which indicates a need for personalization, for which drivers are willing to pay. Furthermore, traffic information and dynamic route guidance lead to significant travel time savings for individual travellers in case of accidents. The exit choice model developed proved to be a valid model, which is a promising start to apply new en-route choice models in various cases. It is recommended to further explore the exit choice model, for example by incorporating habit and learning behaviour or other decision rules. The AIDA RCS could aid such research by providing the necessary environment for data acquisition on a wide range of route choice studies.

Samenvatting

In de afgelopen eeuwen is de mens zich wezenlijk anders gaan verplaatsen. De auto is voor het woon-werk verkeer het favoriete vervoermiddel geworden. In toenemende mate is het wegennet niet in staat om voldoende capaciteit te bieden voor de vraag naar asfalt dat nodig is voor al deze auto's. Dit gebeurt met name in de ochtend- en avondspits in dichtbevolkte stedelijke gebieden, zoals de Randstad. De gevolgen hiervan zijn files en vertragingen. Dit zorgt voor significante kosten voor de samenleving en de economie; direct doordat men tijd niet anders kan besteden en indirect doordat de onbetrouwbaarheid van het verkeerssysteem toeneemt met als gevolg dat een reiziger eerder dan noodzakelijk vertrekt. Zonder congestie is het resultaat van vraag en aanbod een verkeersstroom op het wegennetwerk, met een voorspelbare uitkomst in termen van reistijd, kosten, etc. Een rit is onbetrouwbaar als de verkeersstromen leiden tot onverwachte condities voor de reiziger, met een negatieve beleving en ervaring tot gevolg. Betrouwbaarheid is dus gerelateerd aan de perceptie, ervaring en bekendheid van de individuele reiziger en kan worden opgesplitst in variatie en voorspelbaarheid. Veel variatie en een lage voorspelbaarheid zorgen voor een grote onbetrouwbaarheid. Automobilisten kunnen hun gedrag en keuzes aanpassen om rekening te houden met onbetrouwbaarheid in hun reizen, bijvoorbeeld door het kiezen van een andere route of het aanpassen van de vertrektijd.

Advanced Traveller Information Systems (ATIS) hebben als doel om reizigers te informeren over de huidige en toekomstige toestand van het wegen- of transportnetwerk, zodat de reiziger zijn verwachtingen kan bijstellen of zijn ervaring kan verbeteren. ATIS hebben in bestaand onderzoek veel aandacht gekregen in relatie tot route keuze, vertrektijdstip keuze en zelfs keuze voor vervoerswijze of bestemming. Veelal is hierbij het uitgangspunt gehanteerd dat informatie gebruikt wordt voor een optimale benutting van het wegennetwerk. Persoonlijke verkeersinformatie is gebaseerd op de individuele voorkeuren van de reiziger, maar neemt ook de context van de huidige reis mee. Deze context kan van groot belang zijn. Indien een reis door een weggebruiker vaak wordt gemaakt, is de behoefte voor verkeersinformatie naar verwachting anders dan voor een unieke reis.

In dit proefschrift is de combinatie van niet-recurrente verkeerssituaties (een combinatie van de context van de reis en de toestand van het wegennetwerk) en verkeersinformatie onderzocht. Daarbij is de volgende doelstelling gehanteerd: ***Het verkrijgen van meer inzicht in de impact van verkeersinformatie op routekeuze gedrag in niet-recurrente verkeerssituaties***.

Verkeerssituaties worden veelal weergegeven vanuit het oogpunt van het wegennetwerk of de beheerder hiervan. Files, als gevolg van een vraag die het aanbod overstijgt, worden vervolgens vaak aan de hand van hun specifieke oorzaak benoemd. Oorzaken zoals slecht weer kunnen echter variatie veroorzaken in het optreden van files, wat dit soort files niet-recurrent maakt. Dit oogpunt neemt echter de reizigers die deze file ervaren niet mee in de beschouwing. Deze ervaring speelt naar verwachting een belangrijke rol om te bepalen of een verkeerssituatie wel of niet recurrent is. In dit onderzoek is een kader ontwikkeld dat beide kanten beschouwt. Dit kader combineert files, oorzaken en gevolgen hiervan met de relevante kenmerken van de individuele reiziger. Een reiziger heeft een bepaalde mate van bekendheid en ervaring met de reis die gemaakt wordt, in termen van het wegennetwerk en de verkeerssituaties die hierop voorkomen. Gecombineerd met de perceptie hiervan, leidt dit tot een zeker of onzekere verwachting van de betrouwbaarheid voor de te maken reis. Gebaseerd op de verkeersstromen is het mogelijk om belangrijke objectieve kenmerken vast te stellen zoals variatie en betrouwbaarheid van reistijden. De kenmerken variëren van statistische concepten zoals standaardafwijking tot het bepalen van de extra benodigde tijd om een vastgesteld aantal keren op tijd aan te komen. Met behulp van het kader zijn drie verschillende verkeerssituaties geselecteerd voor verder onderzoek: (a) ongevallen, (b) wegwerkzaamheden en (c) grote evenementen.

Deze drie niet-recurrente verkeerssituaties zijn onderzocht in een enquête over de behoeften naar en voorkeuren voor verkeersinformatie van weggebruikers. De enquête had als doel de relatie te onderzoeken tussen situaties en kenmerken van de automobilist aan de ene kant en kenmerken, types en inhoud van verkeersinformatie aan de andere kant. Deze relatie maakt het mogelijk om verkeersinformatie persoonlijk te maken op basis van de individuele kenmerken en de context van de reis. In totaal 598 Nederlandse weggebruikers hebben de enquête op Internet ingevuld.

De analyse van de resultaten van de enquête laat zien dat weggebruikers de meeste voorkeur geven aan advies voor de snelste route, gevolgd door uitgebreide verkeersinformatie en verwachte aankomsttijd. Afhankelijk van de situatie varieerde de voorkeur voor verkeersinformatie, zonder de volgorde van de drie belangrijkste voorkeuren te veranderen. De actuele verkeerssituatie (bijv. een ongeluk), het motief van de automobilist voor de reis (bijv. werk) en de bekendheid van de bestuurder met de verkeerssituaties waren significant van invloed op de voorkeur voor verkeersinformatie. Ook persoonlijke kenmerken waren een belangrijke verklarende factor voor de keuze voor de inhoud van de verkeersinformatie. Mobiliteitsbeleving, leeftijd en het huidige gebruik van verkeersinformatie waren hierin significant van invloed.

Bestuurders gaven de grootste voorkeur aan advies of persoonlijk advies, zodra men bereid was voor de verkeersinformatie te betalen. In totaal was circa 60% van de respondenten bereid om te betalen voor de verkeersinformatie. De overige 40% koos voor het niet ontvangen van verkeersinformatie als de enige gratis mogelijkheid. Om te bepalen hoe bestuurders deze kenmerken waarderen zijn discrete keuze modellen geschat. Zoals verwacht mag worden zijn kosten het belangrijkste kenmerk van verkeersinformatie voor een keuze door de weggebruiker.

Daarna volgt betrouwbaarheid van verkeersinformatie. Het minst belangrijk vond men de tijdigheid. Mobiliteitsbeleving is hier ook weer gebruikt om te onderzoeken in hoeverre persoonlijke kenmerken van invloed zijn op deze keuze en de waardering van de karakteristieken van verkeersinformatie. Hieruit bleek dat de keuzes tussen de diverse mobiliteitsbelevingen verschilt, maar dat de volgorde van de waardering van de karakteristieken van verkeersinformatie niet verandert. Op basis van de resultaten van de enquête kan worden geconcludeerd dat de weggebruikers behoefte hebben aan persoonlijke verkeersinformatie in relatie tot inhoud, type en kenmerken.

Tijdens ongevallen hebben het advies voor de snelste route en de uitgebreide verkeersinformatie (locatie, lengte, oorzaak en verwachte duur van de files) de grootste voorkeur voor inhoud van de verkeersinformatie. In het onderzoek is een eigen routekeuze simulator ontworpen en ontwikkeld (de AIDA RCS, wat staat voor Applications of Integrated Driver Assistance Route Choice Simulator). Deze simulator is ontworpen om ongevallen en de gevolgen op het verkeer, samen met verkeersinformatie en route advies, te kunnen onderzoeken. De AIDA RCS bestaat uit een micro simulatie model voor het modelleren van de verkeersstromen, een server die zorgt voor de communicatie met het model en het opslaan van gegevens en een client die aan de bestuurder een beeld geeft van de komende kruising, samen met een dashboard en verkeersinformatie en/of een route advies. De bestuurder kan de routekeuze bepalen door het kiezen van een van de armen van de kruising. De server slaat al deze keuzes op, samen met de ontvangen adviezen en informatie en data van de verkeersstromen.

De AIDA RCS is gebruikt in een experiment om de impact van verkeersinformatie en route advies op route keuze gedrag tijdens niet-recurrente verkeerssituaties te bepalen. Hiervoor is een combinatie van een stedelijk netwerk, een ochtend reis naar het werk en een ongeval gekozen als de te onderzoeken niet-recurrente verkeerssituatie. Drie herkomst en bestemming paren zijn gebruikt voor het definiëren van een scenario, elk met twee verschillende ongevallen. In totaal hebben 30 mensen deelgenomen aan het experiment, en elke deelnemer heeft 8 ritten gemaakt in één scenario. De gerapporteerde validiteit door de deelnemers van de AIDA RCS en het experiment zijn gebaseerd op vragen na elke reis. Over het algemeen gaf men aan dat het een realistische ervaring was, met realistische reizen. De deelnemers gaven aan dat ze in de werkelijkheid een gelijke routekeuze zouden maken. Deze hoge mate van validiteit komt overeen met de validiteit van andere routekeuze simulators zoals daarover is gepubliceerd.

De resultaten lieten zien dat deelnemers hun routekeuze na het ontvangen van verkeersinformatie aanpasten en verder hun routekeuze hebben "verbeterd" na het ontvangen van route advies. De verwachte concentratie van verkeer als gevolg van verkeersinformatie en route advies werd gevonden, doordat de variatie in de routekeuze tussen de deelnemers afnam. In elk scenario was één van de ongevallen geplaatst op wat de voorkeursroute bleek te zijn, omdat deelnemers duidelijk hun route aanpasten na het tegenkomen van het ongeval of de gevolgen hiervan. De ongevallen veroorzaakten een toename van de reistijd met bijna 10% en een toename van de vertraging van ongeveer 30% voor de ritten waarin geen verkeersinformatie beschikbaar was. Verkeersinformatie en route advies zorgden voor een afname van de reistijd met respectievelijk 13% en 6%. Beiden zorgen ook voor een afname van de variatie in reistijd en vertraging en een afname van de lengte van de gereden routes. De individuele keuzes lieten zien dat de deelnemers in de meeste gevallen de snelste arm van een kruising kozen (88%), wat toenam tot 94% wanneer de deelnemers route advies ontvingen. Ongevallen hadden een significant effect op de keuzes voor de snelste arm, met name in de situaties waar het ongeval op de voorkeursroute was geplaatst.

De individuele keuzes bij elke kruising zijn daarnaast gebruikt als input voor een uitgangskeuzemodel. Dit model is gebaseerd op de aanname dat een bestuurder bij elke afzonderlijke kruising een keuze voor een route maakt, onafhankelijk van de voorgaande keuzes. Dit reduceert het aantal alternatieven per keuzemoment tot twee of drie, afhankelijk van het aantal armen van de kruising. Om zeker te zijn dat de armen van deze kruising wel naar de bestemming leiden, is de snelste route naar de bestemming berekend, gebaseerd op de uitgang. De gegevens van deze routes zijn gebruikt om de alternatieven (de armen) te beschrijven en vormen daarmee de keuzeset in het model. De beslissingsregels in het discrete keuze model zijn gebaseerd op het principe dat mensen kiezen voor het alternatief met het hoogste nut. Het multinomial logit model is gebaseerd op dit principe, gecombineerd met aannames over de statistische kenmerken van de onbekende elementen in het nut van een alternatief. De karakteristieken van de alternatieven in het model zijn de vrije reistijd (gebaseerd op de maximum snelheid), de vertraging, het aantal kruisingen en een dummy variabele die aangaf of de route de loodrechte afstand naar de bestemming verkleinde. Van deze karakteristieken bleek na het schatten van het model de vrije reistijd de belangrijkste, waarbij bij langere vrije reistijd een extra minuut minder van invloed is dan voor een kortere vrije reistijd. Dit is in overeenstemming met het verwachte gedrag ten aanzien van reistijd in routekeuzes. Vertraging werd gevonden als het daaropvolgende belangrijkste attribuut.

Additionele attributen in het keuze model voor de verkeerssituatie en voor verkeersinformatie zorgen voor een significante verbetering van de modelschatting. Het hebben van een advies voor een bepaalde uitgang bleek daarin van groter belang voor de keuze dan de vrije reistijd van de daarbij horende route. De variabele voor ongevallen op de route heeft geen significant effect op de routekeuze, maar laat wel zien dat een route zonder ongevallen door de deelnemers sneller wordt gekozen. Aanpassing van de beslisregels in het keuze model, door minder restrictieve aannames te doen dan in het multinomiale logit model, laten geen andere bevindingen zien ten aanzien van het keuzegedrag. Alleen in de vergelijking van de drie scenario's onderling is er een significant verschil tussen de geschatte waarden. De specifieke herkomst bestemming relatie kan van invloed zijn op de waardering van de verschillende kenmerken van routes. Deze uitgebreide modellen laten ook zien dat de keuze voor een arm onafhankelijk is van de keuze bij vorige kruisingen en dat de routekeuze in de huidige rit onafhankelijk is van de routekeuze in voorgaande ritten. Daaruit blijkt er in het experiment geen significant leereffect te zijn geweest en is de vermindering in reistijd en vertraging toe te wijzen aan de verkeersinformatie of route advies. Bij het toepassen van het geschatte model op een deel van de dataset wat niet is gebruikt voor het schatten, blijkt dat de voorspellingen van dit geschatte model goed overeenkomen met de daadwerkelijke keuzes.

Dit proefschrift biedt inzicht in de impact van verkeersinformatie op routekeuze gedrag tijdens niet-recurrente verkeerssituaties. Het kader voor niet-recurrente verkeerssituaties biedt hiervoor een breder perspectief dan alleen oorzaken van congestie en kan worden gebruikt voor het vaststellen van de behoefte aan en voorkeuren voor verkeersinformatie. De behoefte en voorkeuren blijken af te hangen van de specifieke verkeerssituatie en de persoonlijke kenmerken van een reiziger. Dit geeft aan dat er een behoefte is om de verkeersinformatie persoonlijk te maken, waarna reizigers bereid zijn om hier voor betalen. Verkeersinformatie en route advies leiden vervolgens tot significante verminderingen van reistijd, voor individuele bestuurders, in situaties met ongevallen. Het uitgangs keuzemodel dat is ontwikkeld is daarbij een valide model gebleken. Het is een eerste toepassing van een nieuw soort routekeuze model om ook tijdens ritten het routekeuze gedrag te kunnen evalueren. Het verdient aanbeveling om dit model verder te ontwikkelen en te kijken naar gewoontegedrag en leergedrag voor dit model, in combinatie met andere beslisregels. De AIDA RCS kan hierbij het hulpmiddel zijn door de omgeving te bieden waarin data voor een dergelijk onderzoek kan worden verzameld.

About the author

Thijs Muizelaar was born in Emmen, The Netherlands on the 26th of January 1980. He obtained his VWO diploma at the O.S.G. De Rietlanden in Lelystad. After that, Thijs studied Civil Engineering and Management at the University of Twente. He choose to study the field of traffic engineering and traffic management, for which he did his internship at ITS Leeds in 2001. In this internship Thijs applied the microsimulation model DRACULA to the city of York to study the effects of an one-way inner ring road. Besides his study, he was treasurer of a committee organising a symposium on multiple land use, and was board member of the local student cycling union and student union. He finished his study with a thesis on the reconstruction of traffic streams on highways using loop-detector data.

After graduating in November 2003, Thijs started his PhD study at the Centre for Transport Studies of the University of Twente, under the supervision of Prof. Bart van Arem. This PhD study was undertaken within the framework of knowledge centre Applications of Integrated Driver Assistance (AIDA). During his PhD Thijs conducted several PR activities, supervised Master students doing their final thesis study, and prepared and presented lectures to Master students on Intelligent Transport Systems (ITS).

Since September 2008 Thijs is employed at Capgemini as a business analyst in the field of mobility. He did several projects for various clients, amongst which Rijkswaterstaat, the executive arm of the Dutch Ministry of Infrastructure and the Environment.

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 100 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Muizelaar, T.J., *Non-recurrent Traffic Situations and Traffic Information: determining preferences and effects on route choice*, T2011/16, December 2011, TRAIL Thesis Series, the Netherlands

Cantarelli, C.C., *Cost Overruns in Large-Scale Transportation Infrastructure Projects: a theoretical and empirical exploration for the Netherlands and Worldwide*, T2011/15, November 2011, TRAIL Thesis Series, the Netherlands

Vlies, A.V. van der, *Rail Transport Risks and Urban Planning: solving deadlock situations between urban planning and rail transport of hazardous materials in the Netherlands*, T2011/14, October 2011, TRAIL Thesis Series, the Netherlands

Pas, J.W.G.M. van der, *Clearing the Road for ISA Implementation? Applying adaptive policymaking for the implementation of intelligent speed adaptation*, T2011/13, October 2011, TRAIL Thesis Series, the Netherlands

Zegeye, S.K., *Model-Based Traffic Control for Sustainable Mobility*, T2011/12, October 2011, TRAIL Thesis Series, the Netherlands

Mhr, T., *Vehicle Routing under Uncertainty*, T2011/11, September 2011, TRAIL Thesis Series, the Netherlands

Pel, A.J., *Transportation Modelling for Large-scale Evacuations*, T2011/10, July 2011, TRAIL Thesis Series, the Netherlands

Zheng, F., *Modelling Urban Travel Times*, T2011/9, July 2011, TRAIL Thesis Series, the Netherlands

Vlassenroot, S.H.M., *The Acceptability of In-vehicle Intelligent Speed Assistance (ISA) Systems: from trial support to public support*, T2011/8, June 2011, TRAIL Thesis Series, the Netherlands

Kroesen, M., *Human Response to Aircraft Noise*, T2011/7, June 2011, TRAIL Thesis Series, the Netherlands

Nielsen, L.K., *Rolling Stock Rescheduling in Passenger Railways: applications in short-term planning and in disruption management*, T2011/6, May 2011, TRAIL Thesis Series, the Netherlands

Qing, O., *New Approach to Fusion of Heterogeneous Traffic Data*, T2011/5, May 2011, TRAIL Thesis Series, the Netherlands

Walta, L., *Getting ADAS on the Road: actors' interactions in Advanced Driver Assistance Systems deployment*, T2011/4, April 2011, TRAIL Thesis Series, the Netherlands

Lin, S., *Efficient Model Predictive Control for Large-Scale Urban Traffic Networks*, T2011/3, April 2011, TRAIL Thesis Series, the Netherlands

Oort, N. van, *Service Reliability and Urban Public Transport Design*, T2011/2, April 2011, TRAIL Thesis Series, the Netherlands

Mahmod, M.K.M., *Using Co-Operative Vehicle-Infrastructure Systems to Reduce Traffic Emissions and Improve Air Quality at Signalized Urban Intersections*, T2011/1, March 2011, TRAIL Thesis Series, the Netherlands

Corman, F., *Real-Time Railway Traffic Management: dispatching in complex, large and busy railway networks*, T2010/14, December 2010, TRAIL Thesis Series, the Netherlands

Kwakkel, J., *The Treatment of Uncertainty in Airport Strategic Planning*, T2010/13, December 2010, TRAIL Thesis Series, the Netherlands

Pang, Y., *Intelligent Belt Conveyor Monitoring and Control*, T2010/12, December 2010, TRAIL Thesis Series, the Netherlands

Kim, N.S., *Intermodal Freight Transport on the Right Track? Environmental and economic performances and their trade-off*, T2010/11, December 2010, TRAIL Thesis Series, the Netherlands

Snelder, M., *Designing Robust Road Networks: a general design method applied to the Netherlands*, T2010/10, December 2010, TRAIL Thesis Series, the Netherlands

Hinsbergen, C.P.IJ. van, *Bayesian Data Assimilation for Improved Modeling of Road Traffic*, T2010/9, November 2010, TRAIL Thesis Series, the Netherlands

Zuurbier, F.S., *Intelligent Route Guidance*, T2010/8, November 2010, TRAIL Thesis Series, the Netherlands

Larco Martinelli, J.A., *Incorporating Worker-Specific Factors in Operations Management Models*, T2010/7, November 2010, TRAIL Thesis Series, the Netherlands

Ham, J.C. van, *Zeehavenontwikkeling in Nederland: naar een beter beleidsvormingsproces*, T2010/6, August 2010, TRAIL Thesis Series, the Netherlands

Boer, E. de, *School Concentration and School Travel*, T2010/5, June 2010, TRAIL Thesis Series, the Netherlands

Berg, M. van den, *Integrated Control of Mixed Traffic Networks using Model Predictive Control*, T2010/4, April 2010, TRAIL Thesis Series, the Netherlands

Top, J. van den, *Modelling Risk Control Measures in Railways*, T2010/3, April 2010, TRAIL Thesis Series, the Netherlands

Craen, S. de, *The X-factor: a longitudinal study of calibration in young novice drivers*, T2010/2, March 2010, TRAIL Thesis Series, the Netherlands