# Effects of Aging on Learning in a Dual-Task Driving Environment
## An Empirical Study

Von der Fakultät 12 - Erziehungswissenschaft und Soziologie -
der Technischen Universität Dortmund
zur Erlangung des akademischen Grades
Doktor der Philosophie
- Dr. phil. -
genehmigte Dissertation.

vorgelegt von
Vera Berthon-Donk

Gutachter:
PD Dr. Gerhard Rinkenauer
Prof. Dr. Michael Falkenstein

Dortmund, im Juli 2013

# Acknowledgements

This thesis was funded by a European Union Marie Curie Initial Training Network (Grant Nr. 238833). Without this funding, this PhD would not have been possible.

I would like to thank my supervisor, PD Dr. Gerhard Rinkenauer for choosing me as a PhD-student for this grant and his continuous support throughout my PhD. Many special thanks go to Dr. Marc Grosjean for his great engagement in this study, his scientific support and his always fair-but-critical look on things.

Great thanks to the student assistants Corinna Schönbeck and Vera Wissing who helped at gathering data, dealing with countless participants, and making sure that studies were always conducted in a correct manner and a reasonable amount of time. Special thanks go out to the various people who provided me with technical support at IfADo, without whom none of these studies would have been possible (in alphabetical order): Tobias Blanke, Ludger Blanke, Hanno Mussmann and Andreas Volgmann.

Special thanks to some (former) employees of IfADo who became friends for life: Katrin Rapp, Sandra Sülzenbrück and Ann-Katrin (Tine) Risse. Thanks a lot for having been there for me, when I either was trapped again in some programming issues, when I did not have the courage anymore to continue or when I just needed a good laugh while enjoying some French cheese and a good red wine.

And last but definitely not least, many thanks to my family. First of all my husband, Jean-Christophe Berthon, for always providing me with support in my career choices, even when that sometimes means living far from each other. My son Lieven, for just being as you are and helping me see what is really important in life. And finally my second child, which will come just after having defended this thesis, for giving me the strength and courage to finish this work before your arrival!

Thanks a lot! Merci beaucoup! Hartelijk bedankt! Vielen Dank!

# Abstract

The purpose of the present thesis was to explore a particular question within the research domain of dual-task driving: Can older people learn how to use an Intelligent Driver Support System? Previous research had provided evidence that driving is a complex task, in which the driver not only receives and analyzes continuous information from the road scene (Jackson, Croft, Kennedy, Owens, & Howard, 2013), but also needs to produce a coordinated motor output, to assure a safe maneuvering of the vehicle (Anstey, Wood, Lord, & Walker, 2005).

Driving is therefore considered multi-tasking and a variety of Intelligent Driver Support Systems (IDSS) have been developed in recent years to alleviate the driver in his task. Research has shown though, that, although those systems become very useful over time, they initially often lead to an increase in overall task complexity (Hancock & Parasuraman, 1992; Vollrath, Schleicher, & Gelau, 2011) due to dual-task interference: The driver has to divide his attention between the driving task and the IDSS leading to loss of performance on either one of each or both tasks. The impact of a secondary task on driving performance is especially visible for older adults.

Due to aging effects, affecting perceptual, cognitive and sensory-motor performance, older adults over 50 experience more difficulties while driving, particularly when additional tasks need to be performed (Hahn, Wild-Wall, & Falkenstein, 2011; Mahr & Mueller, 2011; Merat, Anttila, & Luoma, 2005; Wilschut, 2009). One line of research which has not received a lot of attention so far, is whether practice can improve dual-task performance in driving situations. According to relevant learning theories, such as the Adaptive Control of Thought − Rational (ACT-R) (Anderson, 1982) and the Skill-Knowledge-Rule (SKR) Model (Rasmussen, 1983), learning will lead to automatization, which will free up resources, eventually leading to an improvement in dual-task situations as interference increases. Age effects play a role in learning in dual-task situations as well, especially because at higher age parallel processing becomes more difficult, leading to an increase in interference. Studies examining the effect of learning in dual-task

driving situations lacked experimental control or lacked the inclusion of different age groups, making it difficult to draw any valid conclusions concerning age-related effects of learning in dual-task driving situations.

In this thesis, all experiments took therefore place in a laboratory-controlled environment, to optimize experimental control. We used the Lane Change Test (LCT; Mattes, 2003) as the primary driving task and a visual search task (Treisman & Gelade, 1980) as a secondary task. A first methodological issue that needed clarification was whether the analysis of the LCT could be adapted to individual driving styles allowing a more precise analysis of individual driving styles. Our first experiment therefore aimed at exploring whether the use of a relative calculation method could be used for defining different segments (more difficult versus easier respectively) within the LCT, whether segments based on this relative calculation method were indeed better adapted to individual lane-change behavior and finally whether the use of a relative calculation method for the definition of both segments would more precisely reflect age differences in tracking performance. Our findings show that the use of a relative calculation method to define the start as well as the length (and therewith the end) of a lane-change maneuver, allow the definition of segment windows, taking into account the individual driving style of each participant. Relative windows furthermore turned out to be more precise and more representative of individual lane-change behavior, hence better reflecting age differences as well. This calculation method furthermore allows the creation of two new variables, Reaction Time until Lane Change and Movement Time which both provide valuable insight into lane-change behavior.

Second, the effect of end-of-block feedback (Summary Knowledge of Results; Schmidt, Lange, & Young, 1990) was examined as a method to shift priorities in a dual-task driving environment towards the primary task. In the second experiment it was explored whether driving performance feedback on the LCT in the form of SKR actually helped participants to prioritize the driving task over the secondary task. We examined furthermore whether age had an influence on feedback as a prioritization tool. Finally it was investigated as well whether feedback on the LCT had an effect on learning in dual-task driving conditions. Results in our experiment provided evidence that SKR feedback was a useful tool to prioritize the driving task in a dual-task paradigm including the LCT and a visual search task. This prioritization effect due to feedback was effective for both younger as well as older adults. In this experiment no benefit of feedback on learning effects was found.

The goal of the first learning experiment was to investigate the effect of aging on learning in dual-task driving situations. Our experiment showed that practice had a positive effect on dual-task driving performance for both younger and older adults. Performance differences between age groups remained however and were especially visible in more difficult driving situations or general increases in dual-task complexity (i.e., in a dual-task situation as compared to a single-task situation), indicating that age differences indeed had an effect on learning in dual-task driving situations. Our experiment showed furthermore that acquired skills remain stable over time, even after a retention period.

Finally in the last experiment, a dual-task study was conducted aimed at examining age effects on practice in a dual-task driving environment, but in which the secondary task was of relevance for the driving task. The results showed that a relevant secondary task, which cannot be ignored, had a strong effect especially on driving performance measures, particularly for older adults. Learning had a beneficial effect on both age groups though: Driving performance increased with practice over sessions. Especially older adults benefited from practice, by increasing their capacity to divide their attention between two tasks that were almost equally demanding.

# Zusammenfassung

Ziel dieser Doktorarbeit war es eine präzise Fragestellung zu erforschen im Bereich des Fahrens mit Doppelaufgaben: Können ältere Leute die Benutzung eines Fahrerassistenzsystems erlernen? Frühere Studien haben gezeigt, dass das Fahren eine komplexe Aufgabe ist, in der der Fahrer nicht nur kontinuierlich empfangenen Informationen aus der Fahrumgebung analysieren muss (Jackson et al., 2013), sondern auch eine koordinierte, motorische Ausgabe produzieren muss, damit ein sicheres Manövrieren des Fahrzeugs gewährleistet wird (Anstey et al., 2005).

Das Fahren wird aus diesem Grund als Multitask-Aufgabe betrachtet und eine vielfalt an Fahrerassistenzsystemen wurden in den letzten Jahren entwickelt um den Fahrer diese Aufgabe zu erleichtern. Untersuchungen haben aber gezeigt, dass obwohl diese Systeme mit der Zeit sehr hilfreich werden, sie anfangs oft zu einer Zunahme der allgemeinen Aufgabekomplexität führen (Hancock & Parasuraman, 1992; Vollrath et al., 2011): Der Fahrer muss seine Aufmerksamkeit über sowohl die Fahraufgabe als auch die IDSS verteilen, was zu Leistungverlüste auf eine oder beiden Aufgaben führen kann. Die Auswirkung einer Sekundäraufgabe auf die Fahraufgabe ist vor allem sichtbar bei älteren Fahrern.

Ältere Erwachsenen über 50, haben oft mehr Probleme beim Autofahren wegen altersbedingte Änderungen die die perzeptuelle, kognitive und sensorimotorische Performanz beeinträchtigen. Dies wir vor allem bemerkbar wenn Nebenaufgaben erledigt werden müssen (Hahn et al., 2011; Mahr & Mueller, 2011; Merat et al., 2005; Wilschut, 2009). Bis heute haben sich nicht viele Studien sich damit beschäftigt, wie die allgemeine Doppelaufgabeperformanz in Fahrsituationen verbessert werden kann. Relevante Lerntheorien entsprechend, sowie die Adaptive Control of Thought − Rational (ACT-R) (Anderson, 1982) und das Skill-Knowledge-Rule (SKR) Model (Rasmussen, 1983), wird Übung zu Automatisierung führen, was dazu führt, dass kognitive Ressourcen freigesetzt werden, die dann zu weniger Interferenz in der Doppelaufgabe und eine Erhöhung der

Performanz führen können. Bisherige Studien die den Effekt von Lernen auf Fahrsituationen mit Doppelaufgabe erforscht hatten, konnten mangelnde Versuchskontrolle vorgeworfen werden. Ausserdem wurden nicht immer mehrere Altersgruppen getestet, was eine valide Schlussfolgerung mit Bezug auf Alterseffekte und Lernen in Doppelaufgabesituationen schwierig gemacht hat.

Alle Experimente in dieser Doktorarbeit haben aus diesem Grund stattgefunden in einer kontrollieren Laborumgebung, damit Versuchskontrolle optimiert werden konnte. Wir benutzten den Lane Change Test (LCT; Mattes, 2003) als Primäraufgabe und eine visuelle Suchaufgabe (Treisman & Gelade, 1980) als Sekundäraufgabe. Eine erste methodologische Untersuchung beschäftigte sich damit wie die Analyse der LCT angepasst werden konnte, damit individuelle Fahrstile besser berücksichtigt werden können. Unser Experiment diente dazu zu definieren ob eine relative Kalkulationsmethode benutzt werden konnte für die Definition von unterschiedlichen Fahrsegmenten (schwierige Fahrsegmente versus einfachere Fahrsegmente) in der LCT. Es wurde auch untersucht, ob Segmente basiert auf diese relative Kalkulationsmethode, tatsächlich besser dem individuellen Fahrverhalten angepasst waren und ob Altersunterschiede in der Spurhalteleistung mit dieses relativen Segmenten besser reflektiert werden konnten. Unsere Ergebnisse zeigen, dass die Nutzung einer relativen Kalkulationsmethode zum Definieren des Starts und Länge (und damit das Ende) eines Spurwechselmanövers, es erlaub, Segmente zu definieren, die die individuelle Fahrstil des Probanden berücksichtigt. Relative Fenster waren dazu auch präziser und representativer für individuelles Spurwechselverhalten, und haben aus diesem Grund Altersunterschiede besser reflektiert. Ausserdem erlaubt diese Methodik die Berechnung von zwei neue Variabeln, Reaction Time until Lane Change und Movement Time. Diese Variabeln liefern wertvolle Einsichte in Spurwechselverhalten.

Zweitens haben wir End-of-Block Feedback (Summary Knowledge of Results; Schmidt et al., 1990) als eine Methode, um die Priorität in einer Dual-Task Fahrumgebung auf die Fahraufgabe zu verschieben, erprobt. In unserem zweiten Experiment haben wir untersucht ob end-of-block-Feedback über die Fahrperformanz in der LCT dazu führt dass Probanden effektiv die Fahraufgabe priorisieren. Wir überprüften desweiteren auch ob Alter einen Einfluss hatte auf Feedback als Priorisierungswerkzeug. Unsere Resultate zeichen, dass SKR Feedback ein nützliches Werkzeug sein kann, um die Fahraufgabe zu priorisieren in einer Dual-Task Paradigmus wo die LCT und eine visuelle Suchaufgabe kombiniert werden. Dieses Priorisierungseffekt durch Feedback war effektiv für sowohl jüngere als auch ältere Erwachsenen. In diesem Experiment fanden wir jedoch keinen Hinweis dafür, dass Feedback sich positiv auswirkt aufs Lernen.

Ziel unseres erstes Lernexperiment war es, den Effekt von Alter auf die Lernfähigkeit in Dual-Task Fahrsituationen zu untersuchen. Die Ergebnisse haben gezeigt, dass Übung ein positives Effekt hatte auf Fahrergebnisse für sowohl jüngere als auch ältere Fahrer. Performanzunterschiede zwischen Altersgruppen blieben jedoch erhalten und wurden vor allem sichtbar in schwierigere Fahrsituationen oder Steigungen in der Aufgabekomplexität (z.B. in einer Doppelaufgabe in vergleich zu einer Einzelaufgabe) was zeigt, dass altersunterschiede tatsächlich einen Effekt auf Lernen in Fahrsituationen mit Doppelaufgabe hatten. Unser Experiment hat auch gezeigt, dass das Erlerntes stabil ist über Zeit, selbt nach einer Periode ohne Übung.

Im letzten Experiment, haben wir eine Doppelaufgabe geteststt mit einem Sekundäraufgabe, die relevant war für die Fahraufgabe, mit dem Ziel Alterseffekte auf die Auswirkung von Lernen in einem Dual-Task Fahrumgebung zu untersuchen. Die Resultaten zeigten, dass eine relevante Sekundäraufgabe, die nicht ignoriert werden kann, ein starkes Effekt vor allem auf Fahrperformanzdaten hatte, besonders für ältere Leute. Übung hatte jedoch einen vorteilhaften Effekt auf beide Altersgruppen: Fahrperformanz stieg an mit Übung über Sessions. Vor allem ältere Erwachsenen profitierten von Übung, dadurch dass die Kapazität Ihrer Aufmerksamkeit über zwei Aufgaben zu verteilen zunahm.

# Contents

**1 General Introduction 1**

1.1 Driving: A Complex Task . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.1.1 Intelligent Driver Support Systems (IDSS) . . . . . . . . . . . . . 3

1.1.2 Dual-Task Performance . . . . . . . . . . . . . . . . . . . . . . . . 6

1.2 Aging Effects . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

1.2.1 Effects of Aging on Dual-Task Performance . . . . . . . . . . . . . 14

1.3 Learning in Dual-Task Situations . . . . . . . . . . . . . . . . . . . . . . 16

1.3.1 Retention Effects . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

1.3.2 Theoretical Concepts of Learning . . . . . . . . . . . . . . . . . . 19

1.3.3 The Importance of Feedback . . . . . . . . . . . . . . . . . . . . . 23

1.3.4 Effects of Aging on Learning in Dual-Task Situations . . . . . . . 25

1.4 Structure and Aims of This Thesis . . . . . . . . . . . . . . . . . . . . . 27

**2 General Method 30**

2.1 Participants . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30

2.2 Tasks and Materials . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30

2.2.1 Lane Change Test (LCT) . . . . . . . . . . . . . . . . . . . . . . . 30

2.2.2 Visual Search Task (VST) . . . . . . . . . . . . . . . . . . . . . . 33

2.2.3 Dual Task . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35

2.2.4 Standardized Tests . . . . . . . . . . . . . . . . . . . . . . . . . . 38

2.3 Design and Procedure . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41

2.4 Data Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44

2.4.1 LCT . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44

2.4.2 VST . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50

2.4.3 Dual Task . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51

2.4.4 Subjective Measures . . . . . . . . . . . . . . . . . . . . . . . . . 52

**3 Experiment 1: Methodological Study 53**
3.1 Introduction . . . . . 53
3.1.1 Research Questions and Hypotheses . . . . . 57
3.2 Method . . . . . 58
3.2.1 Participants . . . . . 58
3.2.2 Procedure . . . . . 58
3.2.3 Data Analysis . . . . . 58
3.3 Results . . . . . 60
3.4 Discussion . . . . . 69
3.5 Conclusion . . . . . 70

**4 Experiment 2: Effects of Feedback on Dual-Task Performance 72**
4.1 Introduction . . . . . 72
4.1.1 Research Questions and Hypotheses . . . . . 73
4.2 Method . . . . . 76
4.2.1 Participants . . . . . 76
4.2.2 Design and Procedure . . . . . 76
4.2.3 Data Analysis . . . . . 77
4.3 Results . . . . . 80
4.3.1 General Effects . . . . . 80
4.3.2 First Research Question . . . . . 82
4.3.3 Second Research Question . . . . . 89
4.3.4 Third Research Question . . . . . 91
4.4 Discussion . . . . . 94
4.5 Conclusion . . . . . 97

**5 Experiment 3: Effects of Aging on Learning in Dual-Task Driving Situations 98**
5.1 Introduction . . . . . 98
5.1.1 Research Questions and Hypotheses . . . . . 98
5.2 Method . . . . . 100
5.2.1 Participants . . . . . 100
5.2.2 Procedure . . . . . 101
5.2.3 Data Analysis . . . . . 101

- 5.3 Results . . . . . 104
  - 5.3.1 General Effects . . . . . 105
  - 5.3.2 First Research Question . . . . . 108
  - 5.3.3 Second Research Question . . . . . 120
- 5.4 Discussion . . . . . 126
- 5.5 Conclusion . . . . . 130

**6 Experiment 4: Aging and Learning in Relevant Dual-Task Driving Situations 132**

- 6.1 Introduction . . . . . 132
  - 6.1.1 Research Questions and Hypotheses . . . . . 134
- 6.2 Method . . . . . 135
  - 6.2.1 Participants . . . . . 135
  - 6.2.2 Procedure . . . . . 136
  - 6.2.3 Data Analysis . . . . . 138
- 6.3 Results . . . . . 139
  - 6.3.1 General Effects . . . . . 140
  - 6.3.2 First Research Question . . . . . 144
  - 6.3.3 Second Research Question . . . . . 146
  - 6.3.4 Third Research Question . . . . . 149
  - 6.3.5 Fourth Research Question . . . . . 150
- 6.4 Discussion . . . . . 153
- 6.5 Conclusion . . . . . 157

**7 General Discussion 159**

- 7.1 Summary of Findings and Interpretations . . . . . 159
  - 7.1.1 Experiment 1: Methodological Study . . . . . 161
  - 7.1.2 Experiment 2: Effect of Feedback on Dual-Task Performance . . . 162
  - 7.1.3 Experiment 3: The Effect of Aging on Learning in Dual-Task Driving Situations . . . . . 164
  - 7.1.4 Experiment 4: The Effect of Aging on Learning in Relevant Dual-Task Driving Situations . . . . . 167
- 7.2 Limitations of This Study . . . . . 168
- 7.3 Practical Relevance . . . . . 170

7.4 Directions for Future Research . . . . . 173
7.5 Conclusion . . . . . 176

**References** **178**

**List of figures** **192**

**A Appendix 1: Results Experiment 1** **i**
A.1 Comparing Single versus Dual-Task Data . . . . . i

**B Appendix 2: Results Experiment 2** **xii**
B.1 Comparing Single versus Dual-Task Data . . . . . xii
B.1.1 Driving Measures . . . . . xii
B.1.2 Visual Search Measures . . . . . xxi
B.1.3 Subjective Measures . . . . . xxiii
B.2 Analyzing Dual-Task Data in Detail . . . . . xxvi
B.2.1 Driving Measures . . . . . xxvi
B.2.2 Visual Search Measures . . . . . xxx
B.2.3 Subjective Measures . . . . . xxxii

**C Results Experiment 3** **xxxiii**
C.1 Comparing Single versus Dual-Task Data . . . . . xxxiii
C.1.1 Driving Measures . . . . . xxxiii
C.1.2 Visual Search Measures . . . . . xlii
C.2 Subjective Measures . . . . . xlvi
C.3 Analyzing Dual-Task Data in Detail . . . . . xlix
C.3.1 Driving Measures . . . . . xlix
C.3.2 Visual-Search Measures . . . . . lii
C.3.3 Dual-Task Measures . . . . . lvi

**D Appendix 4: Results Experiment 4** **lxiii**
D.1 Comparing Task Types: Single-Task, Dual-Task and Dual-R . . . . . lxiv
D.1.1 Driving Measures . . . . . lxiv
D.1.2 Visual Search Measures . . . . . lxviii
D.1.3 Subjective Measures . . . . . lxxviii

D.2 Analyzing Dual-Task Data in Detail . . . . . . . . . . . . . . . . . . . . . . . lxxxv
D.2.1 Driving Measures . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . lxxxv

# 1 General Introduction

## 1.1 Driving: A Complex Task

Driving is a complex task, in which the driver receives continuous information from the road scene. He has to analyze and react to this information accordingly, while taking into account traffic systems, driving regulations, conditions of the vehicle, application of road rules and previous driving experiences (Jackson et al., 2013). On the other hand, a coordinated motor output needs to be produced to assure lateral and longitudinal control, quickly react to potential collision risks and interact with other road users, to assure a safe maneuvering of the vehicle (Anstey et al., 2005). Taking into account the complexity of the task, driving by its very nature should therefore be considered multi-tasking.

In fact, the driving task is constituted of numerous subtasks, operating on different levels of priority and complexity. Michon (1985), based on the work by Allen, Lunenfeld and Alexander (1971), argues that the driving task can be represented by the Three-Level Task Hierarchy Model, which defines the driving task as an hierarchical structure of subtasks. As such, the most basic tasks are those that take place at a so-called operational (or control) level. This level refers to basic vehicle control activities such as pushing the gas or brake pedal, steering, changing gears and blinking. Higher up in the hierarchy, tasks take place on the maneuvering level (or tactical level). This level refers to the interaction with the direct road environment to negotiate common driving situations like for example interacting with other road users, respecting safety distances, taking turns and respecting traffic signs. Finally, the strategic level represents higher-order tasks like trip planning, navigating to reach certain destinations in the most efficient way and avoiding traffic jams by choosing alternative routes or alternative times of departure.

What makes this model interesting for this thesis is the time range inherent to each level (Brouwer, Withaar, Tant, & van Zomeren, 2002). Decisions/actions on the operational level are immediate reactions to (potentially dangerous) traffic situations and operate in the millisecond to second range (e.g., braking because of a child running onto the road). Time pressure on this level is high, as the driver is left with limited time to avoid or deal with demanding situations. Time-pressure is less on the tactical level where decisions (e.g., increasing the time-headway to a leading vehicle) are not immediate reactions to potentially dangerous traffic situations, but reactions to cues or contexts that might predict danger. Decisions and actions on the tactical level generally operate in the second to minute range. They can strongly reduce the probability of ending up in time-pressured operational level tasks (e.g., by respecting a safe distance to the car in front, no emergency braking will be needed when the car in front brakes; Dotzauer, Berthon-Donk, Beggiato, Haupt, & Piccinini, in press). Drivers can also make adjustments on the strategic level even before starting a trip (e.g., avoiding rush hours or routes with high traffic density; Brouwer et al., 2002). This strategic decisions influence task accomplishments on the operational level and minimizes the probability of decision making under time pressure (Dotzauer et al., in press).

When looking at Michon's model (1985) and considering the importance of fast decision making at some levels, it is easy to understand that driving assistance might be useful. First, if a system helps the driver at the operational level, for example by preparing the brakes for a strong braking maneuver, crashes can potentially be avoided or attenuated. A system providing assistance at the maneuvering level, for example by assuring a safe distance to a car in front, can assist the driver in avoiding potential actions or decisions on the operational level. By doing this, time-pressure will be reduced and driving comfort will increase. Finally, on a strategic level, a system providing on-route information such as actual traffic conditions or road works, can help the driver take strategic decisions (e.g., to change of route to avoid a traffic jam). Decisions like these will not only augment the comfort of the drive, and therewith the serenity of the driver, but also avoid getting into critical situations (such as following a car closely e.g., in traffic jams). All this will potentially lead to safer driving. In the section that follows, Intelligent Driver Support Systems (IDSS) and their classification will be considered more into detail, taking into account Michon's (1985) model.

### 1.1.1 Intelligent Driver Support Systems (IDSS)

To assist drivers in safely maneuvering the car while performing numerous tasks and subtasks on different hierarchical levels, a variety of Intelligent Driver Support Systems (IDSS), such as navigation devices, lane-keeping and collision-avoidance assistants have been developed (for more examples see Hummel, Kühn, Bende, & Lang, 2011). Mahr and Müller (2011, p. 116) define IDSS as "a (still incomplete) collection of systems and subsystems towards a fully automated highway system, such as autonomous cars". According to this definition, IDSS have the advantage of increasing safety by reducing or even eliminating driver error (for an example on younger drivers see Caird, Chisholm, & Lockhart, 2008) and enhance efficiency (e.g., through automatization more vehicles can be accommodated and driving will always be highly efficient independent of weather and environmental conditions; Brookhuis, De Waard, & Janssen, 2001).

Although technically speaking, research with IDSS is taking great leaps forwards towards autonomous cars on fully automated highways, research has shown that IDSS are not always beneficial for the human being dealing with such systems. Different studies on Adaptive Cruise Control (ACC), one example of IDSS, have shown negative behavioral adaptation when driving with the system, resulting in higher speed, smaller minimum time headways and larger brake forces (Hoedemaeker & Brookhuis, 1998) as well as an increase in response time to a hazard detection task (Rudin-Brown & Parker, 2004), delayed driver reactions in critical situations (Vollrath et al., 2011) and over-confidence in the system resulting in potentially risky behavior (Wilde, 1989). A general observed effect with IDSS use, and of main importance for the theoretical framework of this thesis, is the fact that reductions in the driver's level of attention (Brookhuis et al., 2001) as well as shifts in attention away from the driving scene (Gruendl, 2005; Ranney, 2008) are observed.

#### IDSS Classification

Different types of IDSS exist varying in their function and the amount of assistance they provide. Figure 1.1 represents an IDSS road map (Heide & Henning, 2006). The x-axis represents time, the y-axis represents the level of automation. Represented are different IDSS both in assisted driving (i.e., the driver is still in control) as well as in autonomous

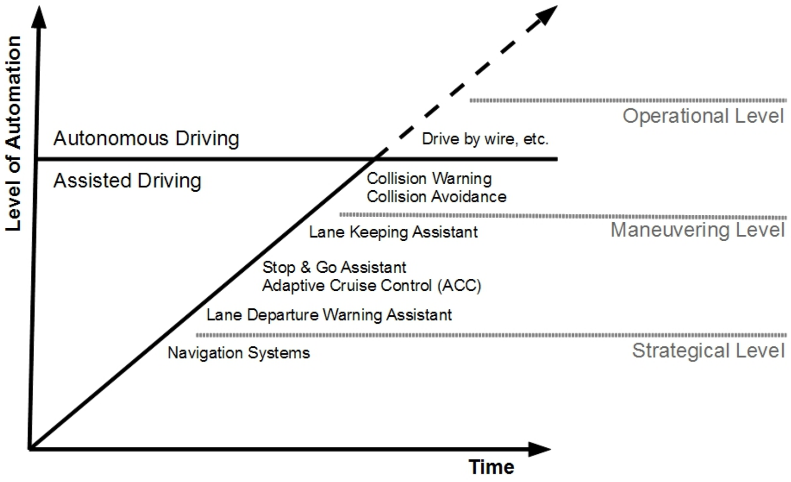


Figure 1.1: Roadmap representing the development of IDSS over time and their increase in complexity (this figure was adapted from Heide & Henning, 2006).

driving (i.e., where the car drives autonomously, without intervention of the driver).

Navigation systems and Lane Departure Warning systems have been on the market for a long time. Those systems are informative. If needed, they warn the driver that he needs to intervene (e.g., when swaying off the road), but they do not actively intervene. Stop & Go Assistants (i.e., systems which overtake car-following situations at low speed, e.g., in stop-and-go situations), Adaptive Cruise Control (ACC; a system which automatically keeps a safe distance to the car in front) and Lane Keeping Assistants (i.e., a system which automatically performs corrective steering-wheel movements when a deviation from a lane is detected) are systems which actively intervene to assure safe driving. Finally, Collision Warning and Collision Avoidance Systems are systems which, based on constant monitoring and interpretation of the traffic environment, can actively take over control to avert potential danger. Today, car-to-car communication, as well as car-to-infrastructure communication (e.g., a traffic light sending information about its status to the car) allow the first experimentations with autonomous driving, that is without intervention of the driver. One example is the "Google Driverless Car" which has driven an important number of kilometers without causing any accidents, respecting the speed limit and traffic rules due to an important number of integrated databases and sensors (Markoff, 2013).

Another interesting development which can be seen in Figure 1.1, is that as time goes by, IDSS are more and more capable of providing assistance at the operational level. If at the beginning, IDSS provided assistance at the strategic level (e.g., navigation systems), with the development of modern technology, more and more assistance is provided on the maneuvering level (eg., Adaptive Cruise Control, Lane-Keeping Assistant). Nowadays and in the near future, systems that assist on the operational level (e.g., Collision Avoidance Systems) help drivers react quickly to dangerous traffic situations, therewith potentially avoiding accidents.

A final consideration when implementing IDSS is to what extent the driver needs to pay attention to them to assure safe driving. Some systems are purely informative, providing information at a strategic level (e.g., a navigation system). If their instructions are ignored by the driver, some inconvenience may occur (e.g., taking a wrong turn or adding a number of kilometers to a planned route), but at no point in time, safe driving will be at risk. However, other IDSS ask for active interventions or decisions of the driver on an operative level (e.g., an ACC without emergency brake system, will warn the driver to brake strongly when the distance to the car in front becomes too small for the system to handle). Ignoring those instructions by the driver might lead to dangerous situations.

Taken together, this overview of IDSS suggests that IDSS can be very useful for the driver, increasing safety when they provide assistance at the time-critical operational level, and increasing comfort when they provide assistance and information at the maneuvering and strategic level. One question of importance for this thesis and which has not received a large amount of attention so far is whether IDSS relevant for the driving task have a stronger effect on driving performance than IDSS without direct relevance for the driving task (i.e., that are of an informative nature only). It is expected that

IDSS relevant for the driving task (i.e., of importance on an operational level) will have a stronger impact on driving performance than IDSS that are of no direct relevance for the driving task.

Although most IDSS become very useful over time, they often lead to an initial increase in overall task complexity (Hancock & Parasuraman, 1992; Vollrath et al., 2011). This increase in complexity is due to dual-task interference, that is, the driver has to divide his attention between both the driving task and a secondary task, which potentially ask for the same resources at the same time (Berthon-Donk, Grosjean, & Rinkenauer, in press). As mentioned earlier, due to its complexity, driving alone can be considered multi-tasking. For the rest of this document, when dual-tasking is addressed, the driving task with an added complementary task is meant, such as driving with an IDSS. This added secondary task, depending on different factors, will have a certain impact on the driving task. In later sections, it will be reviewed how aging and learning can alter the impact from the secondary task on the primary task.

### 1.1.2 Dual-Task Performance

From a theoretical point of view (Pashler, 1994), a classical dual-task paradigm is a situation in which participants are required to perform two tasks simultaneously. To assess the potential costs this may incur, dual-task performance is compared to the respective single-task performances. If a person is able to perform two tasks simultaneously without a drop in performance in either task, then attention is assumed to have been successfully divided. However, if performance drops in either task, then one speaks of dual-task interference. When that occurs, it is assumed that both tasks compete for the same processes or processing resources, thereby potentially limiting attention to one task at a time (for examples, see Olive, 2004).

#### Models of Dual-Task Interference

Several theoretical frameworks have been developed to explain dual-task interference and the presumed failures of divided attention. For example, *Central Information Bot-*

*tleneck Theory* (Pashler, 1984; Ruthruff, Hazeltine, & Remington, 2006) assumes that structural constraints lead to a central bottleneck that only allows us to fully process one task at a time. In particular, certain processes for the second task, such as selecting an appropriate response, cannot start until the same process has been completed for the first task. As a result, one of the two tasks (generally the secondary task) will be performed more poorly (e.g., slower) than in single-task conditions. Along different lines, the *Attentional Resource Theory* proposed by Kahneman (1973) states that dual-task interference is caused by competing demands for attentional resources. Thus, the more tasks that have to be handled simultaneously the less attentional resources will be available for each of them. Kahneman refers to this interference as capacitive interference.

Wickens (1984, 2002) proposed a *Four-Dimensional Multiple-Resource Model* of particular interest for this thesis. His model is an extension to Kahneman's resource theory, that can account for distraction (interference) effects in dual- or multi-task driving situations. According to this model there are four categorical dimensions, each containing different discrete "resource levels". The four dimensions of the model are presented in Figure 1.2 and described below. The basic assumption is that interference will be greater when two (or more) tasks require the same level of a given dimension (e.g., two tasks demanding visual perception).

- *Processing Stages*: This dimension indicates that resources used for perceptual and cognitive activities are different from those needed for the selection and execution of responses. This stage trichotomy is supported by both experimental as well as neuro-anatomical evidence (Van Engelen, 2011; Wickens, 2002, 2008). According to this dimension, there should be more interference between perceptual and cognitive tasks, as they are thought to rely on common resources. For example, visual search on a navigation display coupled with mental rotation (e.g., of map) could lead to interference in a driving situation (Van Engelen, 2011).
- *Perceptual Modalities*: This dimension indicates that auditory perception uses different resources than visual perception. Wickens' model suggests that cross-modal time-sharing (combining a visual with an auditory task) will create less interference than intra-modal time-sharing (e.g., combining a visual task with another

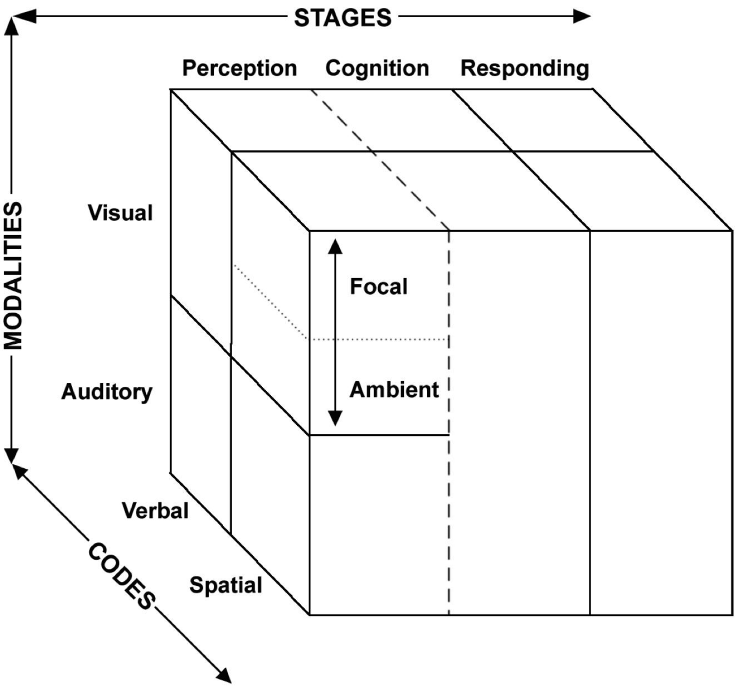


Figure 1.2: Wickens' four-dimensional multiple resource model (this figure was reproduced by permission of Van Engelen, 2011, Figure 3.2, Page 19).

visual task; Wickens, 2002). In a driving environment, an IDSS with a high visual load can easily lead to interference as driving relies heavily on vision as well.

- *Processing Codes*: This dimension indicates that spatial activity uses different resources than does verbal activity. Evidence for this separation comes from studies (Wickens, 2002) showing a relatively efficient time-sharing between manual and vocal responses. Brain research has also shown that this separation can be associated with the functioning of two cerebral hemispheres: The left hemisphere is more specialized in processing nonverbal auditory information, whereas the right hemisphere is more specialized in processing verbal information presented visually. They do this by using resources that are qualitatively different from those of the other hemisphere and that cannot be made available to the other hemisphere (Polson & Friedman, 1988). In driving, which relies greatly on spatial working memory, adding a detailed manual task (e.g., Alm & Nilsson, 1995) might cause more interference than adding a verbal task (e.g., Strayer & Johnston, 2001).

- *Channels of Visual Information*: This dimension is nested within visual spatial resources and refers to two aspects of visual processing: Focal and ambient vision. Focal vision is required for object recognition, visual search and other tasks requiring high visual acuity. For example, searching for information on a navigation display or looking for a specific road sign among others. Ambient vision is used for sensing orientation and ego movement (Horrey, Wickens, & Consalus, 2006), such as for lane keeping or adjusting distance to a leading vehicle. According to Wickens' (Wickens, 2002) framework, focal and ambient vision show efficient time-sharing because they use separate resources.

As Wickens notes himself (2002, 2007), although multi-tasking can theoretically be achieved when the tasks rely on separate resources, certain tasks can demand or attract so much attention that concurrent tasks are ignored altogether. One example he mentions is a study by Strayer and Johnston (2001) who found that drivers became so preoccupied with a cellular phone task, that they would completely neglect aspects of the concurrent driving task, even though the two tasks were not quite structurally related. According to Wickens (2002), two types of interference can thus occur. First, structural interference when two or more tasks require the same processing stages, perceptual modalities, processing codes or visual channels, and second, capacitive interference when multi-tasking leads to conflicts in terms of the overall allocation of attentional resources.

Taking into account the most important models allowing explaining dual-task interference, it is important to review how differential effects, such as aging or learning, can affect dual-task situations. Indeed, healthy aging has an effect on cognitive, sensory-motor as well as perceptual functions, which all influence the human-being's dual-tasking capacity, like for example driving with an IDSS. First, effects of age-related changes on driving performance will be reviewed, before exploring how these changes may potentially affect dual-task performance.

## 1.2 Aging Effects

A main theme of this thesis is the effect of aging on driving performance. Indeed, prognoses of the German Federal Statistical Office (Bundesamt, 2009) foresee an important number of senior citizens in the future. Today, 20 % of all German citizens are over 65 years old. However, it is expected that in 2060, 34 % of all German citizens will be over 65 years old. The combined effect of the demographic situation, improved health-care as well as changes in lifestyle (i.e., driving is an important means of mobility for older people; Chipman, Payne, & McDonough, 1998) will result in an increasing number of older drivers (Lundberg, Hakamies-Blomqvist, Almkvist, & Johansson, 1998) among those senior citizens.

A report of the German Federal Statistical Office (Bundesamt, 2011) analyzing accidents (in Germany for 2011) in which senior citizens ($>$ 65 years old) were involved revealed that in 67 % of the accidents, the older citizens were to blame. According to the report, older adults are often involved in accidents, because "they lose overview in complex situations" as compared to younger traffic participants. Figure 1.3 presents the types of driving errors older drivers make, and how often they led to an accident on German roads in 2011.

As Figure 1.3 suggests, most accidents (18.5%) occur because of errors in "right of way" situations. Driving maneuvers such as turning, driving backwards or approaching follow closely (17,2%). Those accident causes were significantly higher in senior drivers as compared to other age groups and can be explained by the fact that executive functions, needed for the mental construction of an overview of the driving scene (Bundesamt, 2011) decline with age (Adrian, Postal, Moessinger, Rascle, & Charles, 2011). Driving errors such as keeping insufficient distance (8.0%), wrong behavior towards pedestrians (5.8%), not-adapted speed (4.9%), using the wrong side of the road (3.3%), mistakes while overtaking (2.5%) as well as driving under the influence of alcohol (0.9%) occurred less and the observed percentages did not differ from other age groups (Bundesamt, 2011).

Aging is associated with changes in perceptual, cognitive as well as sensory-motor functions (for an extensive resume see Llaneras, Swezey, Brock, Rogers, & Van Cott,

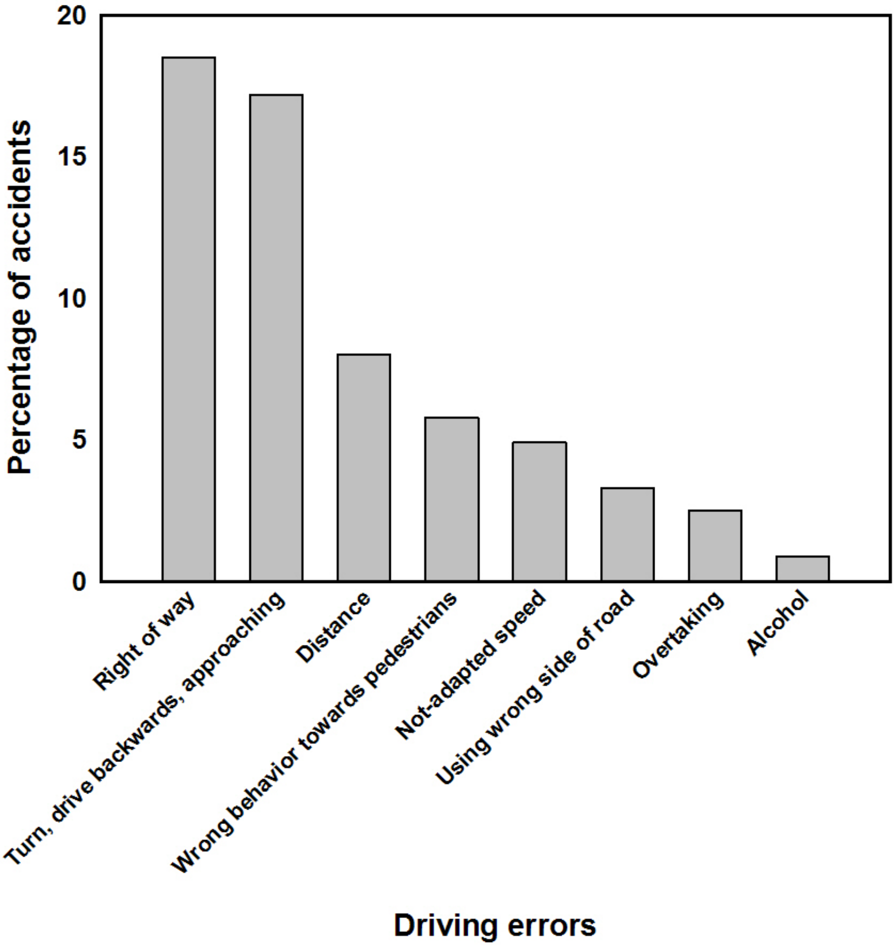


Figure 1.3: Percentage of accidents on German roads in 2011 in which an older driver was at fault, as a function of driving errors (this figure was adapted from Statistisches Bundesamt, 2011, Figure 6, p. 11).

1998). These changes generally cause older adults (over 50 years old) to experience more difficulties while driving, particularly when additional tasks need to be performed, such as paying attention to secondary tasks (Hahn et al., 2011; Merat et al., 2005; Wilschut, 2009), conversing with passengers or on the phone (McCarley et al., 2004) or driving with IDSS (Mahr & Mueller, 2011). In what follows, those changes will be looked at more into detail as well as their potential impact on driving behavior.

### Cognitive Changes

Cognitive changes include significant declines in divided attention, leading for example to worse lane tracking performance (Brouwer, Waterink, Van Wolffelaar, & Rothengatter, 1991; Ponds, Brouwer, & Wolffelaar, 1988; Wild-Wall & Falkenstein, 2010). Selective attention is limited as well at older age (Sweeney, Rosano, Berman, & Luna, 2001) leading to mistakes of omission such as failing to yield (Llaneras et al., 1998). The cognitive decline affects working memory as well (Borella, Carretti, & De Beni, 2008) making it more difficult to discern important from unimportant information (Hahn et al., 2011). Executive functions (Adrian et al., 2011), needed for planning (Allain et al., 2005; Sorel & Pennequin, 2008), problem solving (Diehl, Willis, & Schaie, 1995) and decision making (Henninger, Madden, & Huettel, 2010) are affected by age-related changes as well. In the driving context, older drivers therefore often experience more problems than younger drivers in complex traffic situations (McGwin, Sims, Pulley, & Roseman, 2000), such as an intersection where quick planning, problem solving and decision making are needed (Anstey et al., 2005).

### Perceptual Changes

As most of the information needed for driving is received by the visual system (Owsley & McGwin Jr, 1999), it is not hard to imagine that when those abilities decrease with increasing age, they might lead to poorer driving performance or accidents. According to Llaneras and colleagues (1998) the following perceptual changes might occur with increasing age: Static (i.e., the ability to resolve details of a stationary object) as well as dynamic (i.e., the ability to resolve details of a moving object) visual acuity weakens, leading to difficulties in locating street signs or road boundaries. Contrast sensitivity declines (Owsley, Sekuler, & Siemsen, 1983), which might hinder seeing clearly under poor weather conditions. There is a deterioration in the useful field of view (Edwards et al., 2006) leading to older drivers being less able to detect or process information in the periphery. A degradation in field dependence (i.e., the ability to perceive relevant targets within a number of irrelevant targets) leads to older adults needing longer eye fixations to gather relevant information. Finally there is a decrease in depth perception, which makes it more difficult for older drivers to estimate distances.

### Sensory-Motor Changes

Changes in sensory-motor functions include delayed motor-response speed, leading to increased reaction times (Alm & Nilsson, 1995; Cerella, 1985; Rubichi, Neri, & Nicoletti, 1999; Shanmugaratnam, Kass, & Arruda, 2010), and decreases in flexibility (e.g., Davidse, 2004) and coordination (precision of movement; Poston, Van Gemmert, Barduson, & Stelmach, 2009). Sensory-motor changes might be associated with difficulties controlling lever movements or precise button presses. It should be noted however that modern cars are now more and more equipped with innovative technologies reducing the number of controlled movements (e.g., by the introduction of automatic transmission) or reducing the handling of buttons (e.g., by the introduction of voice command). For this reason, the impact of age-related sensory-motor changes has long been considered minor relative to the influence of perceptive and cognitive age-related changes (Llaneras et al., 1998). With the recent introduction of touch pads in cars however, precise hand coordination while driving is demanded. Older adults might experience more difficulties using this kind of technologies in the car as compared to younger drivers.

An additional difficulty when defining the aging process is that age-related changes vary widely in their onset and amount (Llaneras et al., 1998). Furthermore, for each individual, deteriorations occur at different rates and include different component abilities related to driving. As a result, chronological age alone cannot reliably index the level of competence in basic driving skills and large variations exist (Llaneras et al., 1998).

Based on the overview of age-related changes, it might be interesting to see on what hierarchical level of the driving task (Michon, 1985) older adults adapt to compensate for decline in perceptual, cognitive and sensory-motor capacities. First of all, older adults are known to adapt at the strategic level to cope with complex traffic situations in which they tend to loose the overview of the situation: They avoid rush hours, highly complex intersections, routes with high traffic density (Ball et al., 1998; Brouwer et al., 2002), night-time driving, and driving in poor weather conditions (Ball et al., 1998; Smiley, 1999). In other words: They do adapt to age-related changes by limiting their exposure overall and their high-risk exposure in particular (Ball et al., 1998; Smiley, 1999). On the maneuvering level, tactical changes include driving at lower speed and choosing larger headways to cars in front, to allow for more time to process information and therewith decrease crash risk (Smiley, 1999). Finally at an operational level, few adaptations can be done to cope with difficult situations. Hakamies-Blomqvist, Mynttinen, Backman, and Mikkonen (1999, as cited by Smiley, 1999) found that older drivers managed on average to handle only three basic operational maneuvers (e.g., braking, steering, changing of gear) as compared to four or five maneuvers for middle-aged adults. In other words, on the operational level, due to lacking compensational measures, the older driver is left with limited abilities to decrease accident risk.

Based on the overview above, reviewing the capacities of older adults to adapt to different hierarchical levels of the driving task, it seems obvious that especially IDSS that provide assistance at the operational level (e.g., collision warning and collision avoidance systems) might be useful to increase safe driving for older adults. A question that remains however, is how older adults handle the introduction of IDSS in addition to the driving task. On one hand, due to extended driving experience, they are highly skilled drivers. On the other hand, age-related effects might make the introduction of an IDSS more difficult for older as compared to younger drivers. In what follows, the effect of age-related effects on dual-task performance will be reviewed.

### 1.2.1 Effects of Aging on Dual-Task Performance

A multitude of studies have shown that age-related perceptive, cognitive and sensory-motor changes have an effect on performance in dual-task settings. Interestingly, a meta-analysis on aging and dual-task performance by Riby, Perfect, and Stollery (2004) revealed that age effects are task-dependent: Tasks that rely on controlled processing (i.e., involving an important number of mental resources) showed greater age-related dual-task impairments than tasks that rely on automatic processing (i.e., involving little or no mental resources). For example, Wilschut, Rinkenauer, Brookhuis, and Falkenstein (2008) found differences in how younger (20 to 22 years old) and older adults (50 to 70 years old) switch their attention between driving and a secondary task. The latter consisted of a visual search task that included two difficulty levels, an easier pop-out search and a more difficult conjunction search. Results showed that performance on the driving task as well as the secondary task decreased as compared to (single-task) baseline performance for both age groups. However, the effect was stronger for older adults: In addition to an important drop in driving performance, performance on the secondary task decreased to a point at which it was ignored by the older adults (i.e., they failed to perform the secondary task at all), especially when this task required a more difficult conjunction visual search.

As found by Riby et al. (2004), task-dependency thus seems to play a role in age-dependent dual-task interference, as driving while performing a visual conjunction search requires more controlled processing, which results in a stronger decrease in driving performance for older as compared to younger adults. Wilschut and colleagues (2008) also propose that older adults' allocation of attentional resources is insufficient, resulting in overall performance decreases that go beyond what would be expected from general cognitive slowing due to aging only. Unfortunately the authors fail to explain the reasons for this insufficient allocation of resources. One explanation comes from a study by Hahn, Wild-Wall and Falkenstein (2010). They conducted an experiment aimed at examining age-related differences in stimulus processing in a dual-task driving situation. In accordance with Wilschut and colleagues (2008; 2009), age effects could be observed both in terms of an increase in reaction time on a visual attention task as well as a decrease in tracking performance on a driving-like task. The observed deficit in the secondary visual-attention task was explained by the greater difficulty for older adults (57 to 70 years old) to differentiate between relevant and irrelevant stimuli. According to the authors, this results in a disproportionate amount of attentional resources being allocated to irrelevant stimuli, impeding a proper allocation of resources across tasks.

In light of the studies reviewed above, it seems safe to conclude that aging negatively affects dual-task performance. However, the effects of aging are actually task dependent, in that they mainly arise when tasks involve controlled, as opposed to, automatic processing. Moreover, drops in performance seem not only due to age-inherent cognitive and sensory-motor changes, but also to changes in resource allocation strategies as well, that may rely on people's (decreasing) ability to distinguish relevant from irrelevant information.

One question of importance for this thesis and which has not received a large amount of attention so far is whether age-related effects in dual-task driving situations diminish, or even disappear, with practice. In the following section, learning in dual-task situations will be considered more into detail. We will first define what learning is, how it can be quantified, review models relevant for learning in dual-task situations and discuss the importance of feedback for learning. Lastly, some studies which examined the effect of aging on learning, in general, and on learning in dual-task situations, in particular will be reviewed.

## 1.3 Learning in Dual-Task Situations

The Oxford Dictionary defines learning as "the acquisition of knowledge or skills through experience, practice, or study, or by being taught" ("Learning," 2013). Skill learning is a result of rehearsal or practice. It should be noted that changes in knowledge or skills that are a result of maturation (e.g., a 1-month old baby, being able to track a moving object with both eyes, as opposed to younger babies who cannot do this), the intake of medication or drugs (e.g., leading to changes in perception), structural changes in the brain (e.g., in the case of Parkinson's disease), or due to fatigue (e.g., inducing changes in action sequences) cannot be considered a result of learning (Schermer, 1998, cited by Totzke, Hofmann, Meilinger, Rauch, & Schmidt, 2004). Skill acquisition is a specific form of learning referring to a type of prolonged learning about a family of events that occurs with extended practice. Practice within this framework refers to the effect of repeated task performance and is operationalized by the number of practice trials or the amount of time practicing a specific task (Anglim, 2011). Through many pairings of similar stimuli with particular responses, a person can begin to develop knowledge representations of how to respond in certain situations. These representations can be retrieved more easily and reliably than memories of single events and, as such, skilled behaviors can become routines and even automatic under some conditions (Speelman, 2005). In daily life situations we are constantly exposed to repetition and feedback. This implicit practice increases performance by (partially) automating tasks. This automation of tasks then opens up more capacity for other tasks (Rasmussen, 1983; Shinar, Tractinsky, & Compton, 2005). Within the framework of this thesis, assuming that initial difficulties experienced when driving with an IDSS can be overcome by practice, it is of importance to know whether the results of this learning process persist over time. For this reason, retention effects will be considered as well.

### 1.3.1 Retention Effects

According to Russell and Newell (2007), retention is inherent to learning: Because the construct of learning necessitates retention of what has been learned, learning is assessed through retention tests. The earliest researcher to experimentally examine retention effects was Herman Ebbinghaus (1885a). To observe the process of memory and retention, he constructed a test consisting of several lists of so-called nonsense syllables. He then proceeded to systematically memorizing the items on each list, by reading the first item, saying it to himself, before going to the next item, repeating it to himself, until he reached the end of the list. He assured that the same amount of time was spent on each singular item. After some number of repetitions, Ebbinghaus would attempt to recall the items on the list, until he was able to repeat all items on the list correctly two times in a row. He then waited varying lengths of time before testing himself again. Forgetting turned out to occur most rapidly soon after the end of practice, but the rate of forgetting slowed down as time went by. Ebbinghaus carefully plotted the amount of retention over time, therewith documenting the first forgetting curve. It should be considered that his research (and a lot of research thereafter) was based on learning of facts, whereas for this thesis, retention effects on (motor) skill learning are of relevance. The acquisition of a skill (or any knowledge actually) involves three distinct processes. First, encoding processes are of importance during the practice phase when the learner practices the skills. Then consolidation processes stabilize the acquired skill during the retention period. Finally, during retention tests, retrieval processes allow the learner to reproduce the learned skill. All these processes heavily depend on memory and attention processes (Wagner, 2006). Kantak and Winstein (2012) performed an extensive literature review on (motor) skill acquisition and retention effects. They define a retention test as "an assessment of performance of the same skill under the same conditions that was practiced in the acquisition phase, in order to determine the relative permanence of the level of performance achieved in acquisition" (Kantak & Winstein, 2012, p. 221). In other words: A retention test examines the extent to which a skill is retained by the learner over the retention interval.

A large amount of literature concerning retention effects on skill learning comes from studies in the medical domain: Extensive research has been done on the acquisition and retention of cardiopulmonary resuscitation (CPR) knowledge and skills for adults (Ackermann, 2009; Gombeski, Effron, Ramirez, & Moore, 1982), as well as newborns (Kaczorowski et al., 1998). Most studies have shown that those skills and knowledge deteriorate within weeks of training. In fact, the largest part of skills is forgotten 1 week after the last training session and then forgotten at a slower rate after a 7-8 months interval. These findings are in accordance with initial findings by Ebbinghaus (1885a), who found that forgetting was fastest directly after the end of practice and slowed down thereafter. These findings also raises the question of retention interval length (i.e., the time interval between the end of acquisition and the retention test). Ideally, it is generally assumed to be one that is long enough to provide sufficient time for the processes of memory consolidation to occur, but not so long that there is a loss of performance due to forgetting processes (Russell & Newell, 2007). Kantak and Winstein (2012), in their extensive review of the literature, found that the retention interval is extremely variable across studies, often depending on experimenters' choice and constraints. They categorized the retention tests into either immediate or delayed retention tests. Time intervals for immediate retention tests vary from 10 seconds to several hours following practice. Delayed retention tests are retention tests that occur after a period of at least 24 hours without practice. An extensive review of the literature by these same authors (Kantak & Winstein, 2012) showed furthermore that performance at delayed assessment, was a more reflective measure of the relatively permanent change in the capability for the practiced skill and allowed a more valid inference of how well the learner has encoded, consolidated and retrieved the motor memory.

Typically, studies in the domain of skill learning within the medical domain use retention intervals ranging from 1 week, to several months (e.g., Edelman, Mattos, & Bouwman, 2010), to even one year (e.g., Gombeski et al., 1982). Surprisingly few studies using dual-task driving paradigms including IDSS, have been repeated in time to account for learning effects while driving. Most studies examine single-session learning including a different number of trials (Shinar et al., 2005). Some authors looked at learning effects including plural sessions, ranging from two (Popken, Nilsson, & Krems, 2008) to six sessions (Chisholm, Caird, & Lockhart, 2008). Those sessions took either place on the same day (Popken et al., 2008), on consecutive days (Cooper & Strayer, 2008) or with one to four days in between (Chisholm et al., 2008). To our knowledge, none of those studies looked at the effect of a retention interval on learning to use IDSS with the driving task.

## 1.3.2 Theoretical Concepts of Learning

### Power Law of Practice

One of the methods used to quantify learning has been to find a function that describes how performance changes with practice. Ebbinghaus (1885a) was the first one to describe a learning curve, by carefully documenting the number of repetitions needed to recall a list of nonsense syllables. A well-known example is the *Power Law of Practice* (Newell & Rosenbloom, 1981), which can be expressed with the following equation:

$$T = NP^{-c},$$

where $T$ is the time to complete a task, $P$ is the number of practice trials, and $N$ is performance time on the first trial of the task. The parameter $c$ in this equation is the learning rate. The value of $c$ is usually between 0 and 1, and is preceded by a minus sign to capture the negatively accelerated nature of the learning curve. The closer the value of $c$ is to 1, the faster the learning rate.

Although the power law of practice can be widely applied to all forms of knowledge acquisition, it is based on the assumption that something will slow down the learning process at some point. That is, after a certain amount of practice a so-called learning plateau is often reached, where performance essentially ceases to improve. Such a plateau can be due to a decrease in motivation, a lack of automatization of (partial) actions, a lack of transfer from already existing knowledge to a new task, or an over fixation on parts of the task only, thereby forgetting other parts of the task (Schermer, 1998, , as cited by Totzke et al., 2004).

In order for a skill to become "automatic", that is efficient, unintentional and unconscious (Charlton & Starkey, 2011), it is generally assumed that people go through a number of stages. Two models have tried to define those stages, which will be described in the section that follows.

### Adaptive Control of Thought – Rational (ACT-R)

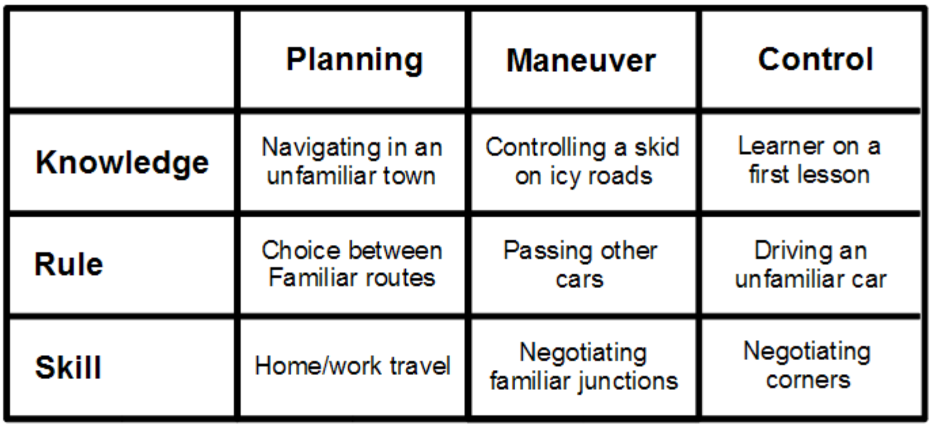
The *Adaptive Control of Thought – Rational* (ACT-R) theory of skill acquisition by Anderson (Anderson, 1982) is a cognitive architecture and has been validated to model human behavior (e.g., while driving; Salvucci, 2006) and learning in a variety of tasks (Kim, Koubek, & Ritter, 2007). The model explains the acquisition of skills via three main stages: A cognitive, an associative and finally an autonomous stage. In the initial, cognitive stage, people solve problems in new domains by applying unspecific problem-solving productions to explicit knowledge they have about this domain. Productions are condition-action pairs that specify that if a certain state occurs in working memory, a particular action should take place. They preexist in peoples' repertoire and are not linked to any specific domain. What is more, they require attention and working memory, making their execution slow, effortful, and relatively inefficient. The second stage of learning, the associative stage, corresponds to the acquisition of a problem-solving routine through knowledge compilation. With practice, productions are progressively adapted to the task at hand, making problem solving more efficient and more reliable. Separate components are combined into one routine. Finally, in the third autonomous stage, routines are triggered in an autonomous manner and do not demand any cognitive resources, thereby freeing up resources to do other activities in parallel.

For skill acquisition to take place according to ACT-R, several memory types exist: Working memory, declarative memory and production memory. Working memory is the part of memory used for interaction with the outside world. It encodes information from the outside world and stores this information in declarative knowledge or performs pattern matching with production rules in procedural knowledge. It also retrieves information from declarative memory or retrieves production rules from procedural memory. Declarative memory is where declarative knowledge is stored and the latter refers to factual information (e.g., Berlin is the capital of Germany). Procedural memory is where production rules are stored. The interaction between these different types of memory allows the development of cognitive strategies resulting in skill acquisition. Examples of such strategies include proceduralisation (the combination of different chunks of explicit knowledge into procedures), composition (a sequence of individual productions is progressively collapsed into a larger task-specific production), and generalization (the application of similar production rules to different situations). All these methods allow freeing up resources in working memory and attention. Based on this theory, one would expect mental load to decrease with practice as actions become more automated. In a dual-task setting, this would imply that more resources should become available for other tasks, hence diminishing the amount of dual-task interference over time.

### Skill-Rule-Knowledge Model (SKR-Model)

Another model, which is not one of skill acquisition per se, but which allows to explain the stages a human being goes through when acquiring practice and experience, is the *Skill-Rule-Knowledge Model* (SKR-Model) (Rasmussen, 1983). According to this model, all human behavior can be broken down into skill-based, rule-based and knowledge-based behavior. It is a hierarchical model that ranks the types of behaviors in terms of how much mental processing they demand. As such knowledge-based behavior occurs when getting into an unfamiliar situation or when learning something new. No "know-how" rules are available and performance is on a conceptual level, which is goal-controlled and knowledge-based. At this level a task is being performed in a highly conscious manner, costs a lot of mental effort and execution of the task is in general slow as each step to master the situation must be defined and tested. Through experience and training, rules and procedures are applied in familiar situations: Behavior becomes rule-based facilitating problem solving and allowing faster decision making. Finally on a skill-based behavioral level, sensory-motor performances become automated and require very little or no conscious control. Both rule- and skill-based behaviors free up cognitive resources.

The SKR-Model and the Three-Level Task Hierarchy Model of Michon (1985) can be combined (Hale, Stoop, & Hommels, 1990) to demonstrate that different behavioral levels are possible at each hierarchical level of the driving task, especially depending on the driving task. The combination of both models is presented in Figure 1.4. The first column represents the three levels of task classification (Michon, 1985) and the first line represents the three levels of behavior (Rasmussen, 1983). Typical examples of each task are presented within the matrix.



| | **Planning** | **Maneuver** | **Control** |
|---|---|---|---|
| **Knowledge** | Navigating in an unfamiliar town | Controlling a skid on icy roads | Learner on a first lesson |
| **Rule** | Choice between Familiar routes | Passing other cars | Driving an unfamiliar car |
| **Skill** | Home/work travel | Negotiating familiar junctions | Negotiating corners |

Figure 1.4: Examples of driving tasks combining Michon's Three-Level Task Hierarchy Model and Rasmussen's Skill-Knowledge-Rule Model (this figure was adapted from Hale et al.; 1990, Figure 1, p. 1383).

For an experienced driver most driving tasks are on and under the diagonal from the upper left to the lower right of the table. The driver tasks encountered by inexperienced drivers are mainly in the quadrant in the upper right corner. With experience, tasks become more automatic and move to a skill-based level. For this reason, drivers operating on a knowledge-based level will operate in a less homogeneous and predictable manner than drivers which operate at a rule- or skill-based level (Hale et al., 1990). It should however be considered, that in some situations, even experienced drivers will show knowledge-based behavior (e.g., when navigating in an unfamiliar area) before acquiring enough experience in this particular situation to move to a rule- or skill-based level again.

The population of interest for this thesis, namely the aging population, is represented at several cells of this table. Older drivers, due to their age, have usually a great amount of driving experience. For this reason they are highly skilled: They know their routes, they know how to negotiate junctions and how to control and operate their car. Furthermore, as mentioned above, they apply strategic tactics to remain at a skill-based level (e.g., by avoiding rush hours, complex intersections, poor weather conditions, etc.; Brouwer et al., 1991; Smiley, 1999). The introduction of IDSS, aimed at helping the driver at his task, can however move them to a knowledge-based level, as those systems not only force drivers how to learn to use those systems, but changes their driving task as well (going from a single- to a dual-task situation). Learning to use those systems can help older drivers being comfortable while assuring a safe drive.

### 1.3.3 The Importance of Feedback

One of the key issues of this thesis is learning. Feedback plays an important role in learning, as it provides motivation (Vollmeyer & Rheinberg, 2000), guidance (Salmoni, Schmidt, & Walter, 1984; Schmidt et al., 1990) and can be used for shifting prioritization between tasks, therewith shifting the focus from one task to another (Levy & Pashler, 2008). In this section we will have a look at different types of feedback and how they can be useful in a driving context.

#### Types of Feedback

Feedback can be immediate, that is, directly after each trial the participant becomes information about his performance. This feedback is often referred to in the literature as Knowledge of Results (KR; Schmidt et al., 1990). It is thought to result in good memory representations of movement, as the participant becomes immediate feedback over his performance. A drawback of this type of feedback however, is that due to the strong informational component of KR, especially in the early stages of learning, KR can strongly guide the participant towards the appropriate movement pattern. As such, the participant might learn to rely upon KR to maintain trial-to-trial performance instead of processing more in-depth task-related information (Salmoni et al., 1984; Lavery, 1962). Another drawback of this type of feedback is the absence of a more stable response pattern, as participants are more likely to alter their response on the next trial after receiving KR, which might lead to higher performance variability (Salmoni et al., 1984).

Feedback can also be presented at the end of each block, known as Summary Knowledge of Results (SKR). This type of feedback helps in-depth learning without causing a lot of variability between trials (Berthon-Donk, Grosjean, & Rinkenauer, 2011; Schmidt et al., 1990). A drawback of this type of feedback however is that, due to the absence of immediate guidance, learning might be slower as the participant does not become immediate feedback over his actions. Both KR as well as SKR should also increase motivation, by providing information on the own performance.

### Feedback in the Driving Context

In a laboratory setting, feedback is often provided as a "value of goodness" (in comparison to other participants or to a previous performance). In real driving contexts, feedback is rarely provided in this manner, but often more implicit in nature (e.g., by using engine noise as immediate feedback to control driver speed; Hellier, Naweed, Walker, Husband, & Edworthy, 2011), aimed at providing immediate information without distracting the driver. Another difference with laboratory settings is the fact that often feedback is provided in real-time or in a concurrent manner, that is at the moment an event occurs. This is often the case with informative IDSS. One example is an Intelligent Speed Adapter (ISA), providing immediate vocal or auditory feedback when the driver passes the current speed limit (which significantly reduces the number of speed violations; Brookhuis & de Waard, 1999). Another example are systems that monitor driver distraction by measuring off-road eye glances and which, thanks to concurrent feedback, dissuade drivers from engaging in distracting activities (Donmez, Boyle, & Lee, 2007, 2008).

Recent studies however have shown, that retrospective or post-hoc feedback after the drive (Zhao & Wu, 2012) can also have advantages over real-time feedback. First, retrospective feedback does not interfere with the driving task. Second, providing post-hoc feedback can help refresh the driver's memory to understand how safe his driving really is (e.g., when the number of traffic light violations is presented after the trip). Third, retrospective feedback which is reported after a trip, leaves more time for a driver to assess and modulate his overall driving behavior without any time or resource constraints (Zhao & Wu, 2012). Indeed, two studies found that retrospective feedback had a positive effect on driving performance both in a safety-critical scenario (Donmez et al., 2008) as well as in a less safety-critical scenario (Zhao & Wu, 2012).

Taken together, this overview suggests that feedback plays an important role in providing guidance and motivation, which are both important prerequisites for learning. Feedback provided in real-time has the advantage of influencing behavior immediately, leading to direct changes in driving behavior. Retrospective feedback has the advantage of being less distracting and allow for an in-depth reflection on driving behavior, leading to positive changes in driving behavior as well. Independent of its nature, feedback thus fulfills its goal to guide and motivate, leading to behavioral adaptation.

### 1.3.4 Effects of Aging on Learning in Dual-Task Situations

If we consider that skill acquisition and dual-tasking rely heavily on memory and attention (Repovs & Baddeley, 2006), which decline at older ages (Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010), learning in dual-task situations should differ between older and younger adults.

At a general level, aging has an effect on cognitive functioning as a whole, resulting in deficient problem-solving skills, poor sustained attention, and an impairment in the generation of reliable goal structures needed for skill acquisition (Peretti, Danion, Gierski, & Grangé, 2002). One reason for a decrease in cognitive functioning at older age is a decline in memory capacities, which has an effect on learning both sensory-motor as well as cognitive tasks. Indeed, Verwey, Abrahamse, Ruitenberg, Jiménez, and De Kleine (2011) found that older adults (55 to 62 years old) make less use of motor chunks when learning movement patterns in a dual-task setting, due to age-related declines in memory functions. This resulted in poorer sensory-motor performance as compared to younger adults.

Memory capacity is also needed to manage and coordinate multiple tasks at the same time. Voelcker-Rehage and Alberts (2007) found that this ability was impaired for older adults (65 to 75 years), even after extended practice on a task including both a sensory-motor as well as a cognitive task. Göthe, Oberauer and Kliegl (2007) found that this decrease in capacity to manage and coordinate multiple tasks at the same time resulted in a qualitative difference in the way younger (16 to 19 years old) and older adults (64 to 77 years old) deal with dual-task requirements. Specifically, most younger adults made a transition from serial to parallel processing with practice. Older adults, however, apparently did not make this transition. The authors explain these findings by age-related changes in the executive system of older adults, which heavily relies on memory processes. This system is needed to assign processing resources or processing time to concurrent tasks, and for scheduling processing steps in a way that minimizes interference (Baddeley, 1996). For example, by holding only one task set in operative mode at any time (serial processing). With practice, the executive system might overrule the serial-processing constraint, allowing parallel processing of two task sets. Older adults' executive systems seem to be more conservative than those of younger adults, in that they cannot overrule the serial-processing constraint and thereby will not achieve perfect time sharing.

Taken together, these studies suggest that age-related declines in memory functions has an effect on learning in dual-task situations, resulting in poorer performance on sensory-motor tasks and difficulties in handling dual-task situations. As this thesis focuses on dual-task situations involving driving, studies that attempted to address these situations will now be considered. Surprisingly, few studies have been done in this domain, despite the recent rise of IDSS and entertainment systems within the car.

Shinar, Tractinsky, and Compton (2005) used a driving simulator to perform a dual-task driving study in which participants simultaneously performed a math operations task or an emotionally involving phone conversation. Their experiment included three driver groups: Young/novice drivers (all, but one, were 18 years old), experienced drivers (30 to 33 years old), and older drivers (from 60 to 71 years old). Practice effects were assessed by looking at how performance changed over 5 sessions that extended over 14 days (with 1 to 4 days in between each session). Although they found an effect of age, with older adults showing poorer performance on all tasks as compared to younger adults, they also found that practice had a positive effect on performance. In particular, age-related differences in driving performance decreased over sessions and by the fifth session all three age groups had nearly identical levels of driving performance. The authors concluded that with sufficient practice, dual-tasking in a driving environment is possible for both younger and older drivers. Unfortunately, these results were questioned by later studies from Cooper and Strayer (2008) and Chisholm, Caird, and Lockhart (2008). Both studies looked at the effects of practice on secondary-task related driver distraction by conducting simulator studies including multiple sessions. Both studies included only one (younger) age group and they both found that, although practice could increase dual-task performance, baseline levels (i.e., single-task performance levels) could never be reached.

A general conclusion from these driving-related dual-task studies is that dual-task performance becomes better with practice. However, special care should be taken when interpreting the results of these studies because the experimental settings might not have allowed for proper experimental control (e.g., as is the case with naturalistic phone calls). Moreover, not all studies included different age groups, making it difficult to draw any general conclusions regarding age-related effects on learning in dual-task driving situations. It thus seems justified to conduct a study in a controlled laboratory environment, aimed at examining the effect of aging on driving and secondary task performance. Of particular interest is the question how performance changes with repeated practice for both age groups. Another question unanswered so far, to our knowledge, is whether older adults develop specific strategies to better cope with demanding dual-task driving situations.

## 1.4 Structure and Aims of This Thesis

The following chapters are organized around different experiments conducted in laboratories and aimed at examining the effect of age-related changes on learning in dual-task driving situations. Previous research had shown that although aimed at improving the driving performance, IDSS often lead to dual-task costs, leading to decreases in performance, especially for older adults (Wilschut, 2009). It thus remained to be explored whether practice could help improve performance for both younger and older adults in a dual-task driving situation. All experiments are based on the same methodological setting, which will be explained in the General Method (Chapter 2). The chapters that follow the General Method (Chapters 3 to 6) are each organized around an experiment aimed at answering one or several precise research questions.

While previous attempts by other authors at (partially) answering similar research questions (e.g., Chisholm et al., 2008; Cooper & Strayer, 2008; Shinar et al., 2005), had been hard to interpret due to methodological issues, experiments in this thesis were conducted in a controlled laboratory environment, using a simulated, simplified driving environment (LCT; Mattes, 2003) as the primary driving task and a visual search task (Treisman & Gelade, 1980) as a secondary task (this experimental setting was based on previous work by Wilschut, 2009). A first methodological issue that needed clarification was whether the analysis of the LCT could be adapted to individual lane-change behavior allowing a more precise analysis of individual driving styles. The first experiment (Chapter 3) therefore first aimed at exploring whether the use of a relative calculation method could be used for defining different segments (more difficult versus easier respectively) within the LCT. A second objective was to examine whether segments based on this relative calculation method were better adapted to individual lane-change behavior. Finally, it was examined whether the use of a relative calculation method for the definition of both segments would more precisely reflect age differences in tracking performance.

A second methodological issue that needed to be clarified was how to make sure participants would prioritize the driving task, despite being in a simulated driving environment. End-of-block feedback (Summary Knowledge of Results (SKR); Schmidt et al., 1990) was provided in addition to explicit instructions to prioritize the driving task over the secondary task. Feedback is an important tool for learning as well, as it provides guidance (Salmoni et al., 1984; Schmidt et al., 1990) and motivation (Vollmeyer & Rheinberg, 2000). In Experiment 2 (Chapter 4) the question whether driving performance feedback on the LCT in the form of SKR actually helped participants to prioritize the driving task over the secondary task was explored. It was furthermore examined whether age had an influence on feedback as a prioritization tool. Finally it was investigated whether feedback on the LCT had an effect on learning in dual-task driving conditions.

The first learning experiment (Chapter 5), aimed at investigating the effect of aging on learning in dual-task driving situations. The second learning experiment, and last experiment of this thesis (Chapter 6), was a dual-task driving study aimed at examining age effects on practice in a dual-task driving environment, but in which the secondary task was of relevance for the driving task. Although some research in that direction had been done (see Seppelt & Wickens, 2003), with this experiment both younger and older adults were pushed to their limits by providing a secondary task which could not be ignored, as it provided direct indication for the driving task. In other words: Ignoring the secondary task would lead to performance loss on the driving task.

In the final chapter (Chapter 7), major findings as well as their interpretations will be reviewed and discussed. Then, general limitations of these studies, practical relevance of this thesis as well as a number of future directions will be discussed.

# 2 General Method

## 2.1 Participants

In total of 130 younger and older individuals participated in 4 experiments. They were all in possession of a valid driver license. Younger participants were mostly college students and recruited by advertisements at the local university and online through social networks. Older participants were recruited out of an existing internal database at the IfADo or by an advertisement in the newspaper. Due to the similarity of methodology in the different experiments conducted, none of the participants took part in more than one experiment. They all had normal or corrected-to-normal vision, reported the absence of neurological impairments and were paid for participation.

## 2.2 Tasks and Materials

Each of the experiments was composed of a standardized driving task (Lane Change Test). This task was either performed on its own or in combination with a visual search task. The combination of the driving task with the visual search task resulted in a dual-task condition.

### 2.2.1 Lane Change Test (LCT)

The driving task consisted of the Lane-Change Test (LCT; Mattes, 2003). The LCT is an automotive tracking task that has the goal to assess changes in performance on lane-change maneuvers while performing additional non-driving related activities (i.e., secondary tasks). The LCT has become an international ISO standard (ISO 26022, 2010) and represents a simplified driving environment. The advantage of the LCT is

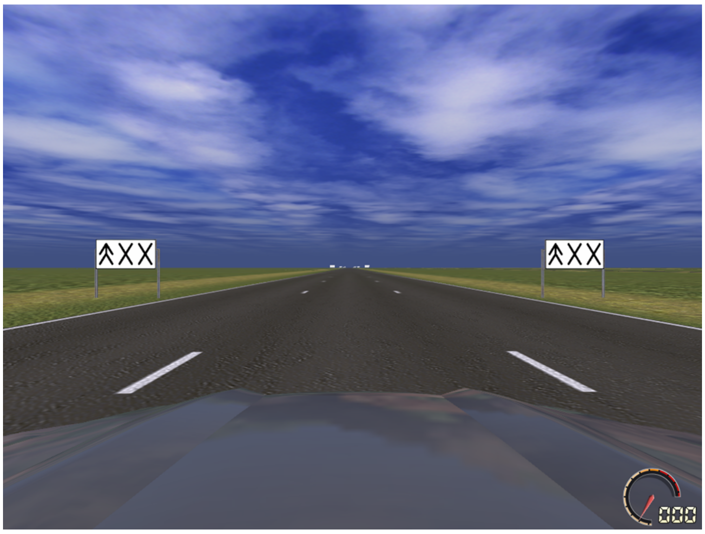


Figure 2.1: Screenshot of the LCT showing the straight 3-lane road without other traffic and with lane-change signs at both sides of the road which, in this case, indicate a lane change to the left of 1 lane.

that it represents a simple, reliable (Benedetto et al., 2011) and sensitive (Bruyas et al., 2008; Harbluk, Mitroi, & Burns, 2009; Maciej & Vollrath, 2009) method to quantitatively estimate secondary task demands, elicited by for example visual-manual or speech-based interfaces (Harbluk, Burns, Lochner, & Trbovich, 2007) in a driving context. The LCT is furthermore sensitive to training effects and can thus be a useful tool for measuring learning in a driving context (Petzoldt, Bär, Ihle, & Krems, 2010; Huemer & Vollrath, 2010). Figure 2.1 presents a screenshot of the LCT. It consists of a straight 3-lane road without other traffic and with lane-change signs at both sides of the road which indicate lane changes to the left or the right of 1 or 2 lanes. Although not visible in the screenshot, each track consists of 18 signs. The mean distance from sign to sign is exponentially distributed, to avoid anticipation by the participant, and ranges from 140 m minimum to 188 m maximum. At a constant speed of 60 km/h, participants get to see a traffic sign approximately every 10 seconds. There are 6 possible lane changes which each occur exactly 3 times on each track.

Figure 2.2 presents the experimental setup. The hardware for the LCT consisted of two IBM compatible PCs running Windows XP®. One was the computer on which the

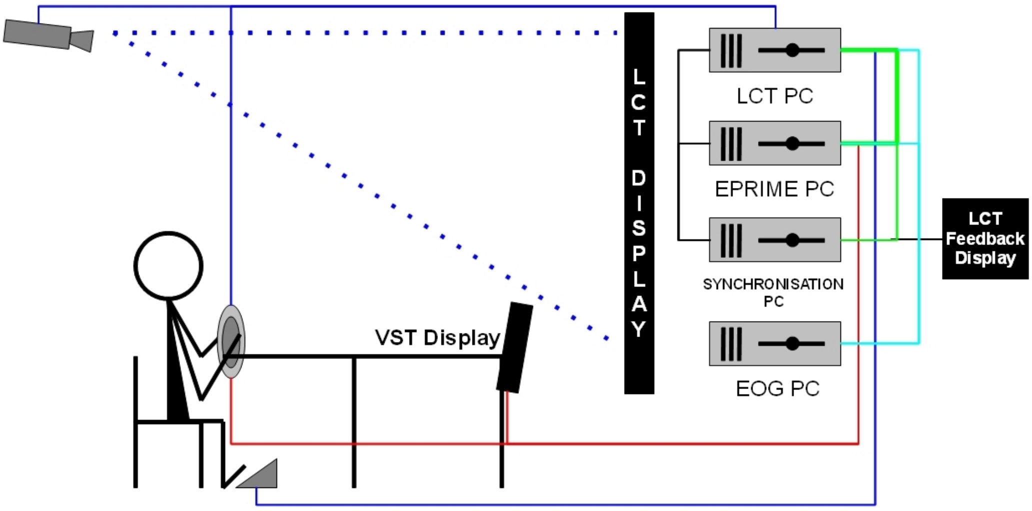


Figure 2.2: Experimental setting representing the position of the LCT projection, the seating position of the participant, the position of the VST display and the function of different computers used (see text for details).

LCT program was running (*LCT PC*). It was connected to foot pedals and a Logitech® steering wheel with 2 standard push buttons on either side on the back. The participant was seated in a height-adjustable office chair in front of a table to which the steering wheel was attached. The viewing distance to the projection screen on which the driving image was presented was approximately 180 cm to 200 cm. The resolution of the projected image was set to 1024 x 768 pixels and subtended a visual angle of 33.4° x 43.4°. The LCT PC recorded the following LCT-variables: Time (ms), x-, and y-coordinates of the virtual vehicle's current position (m), speed (km/h), steering-wheel angle (°) and current track number (1-10).

The second computer (from here on called the *synchronization PC*) provided feedback on the driving task for certain experiments. The LCT feedback display was positioned on the right of the LCT projection, out of the direct visual field, at a visual position of -10 degrees below the horizon of the driving image and about 40 degrees from the middle of the projection screen (the actual position of the LCT feedback display is not shown in Figure 2.2). For data backup reasons, the synchronization PC recorded the same data as the LCT PC.

### 2.2.2 Visual Search Task (VST)

When a secondary task was present, we used a visual search task (VST) as a surrogate In-Vehicle-Information System (IVIS). Wilschut (2009) used a similar experimental setting, including a visual search task in combination with the LCT, to examine the effect of age on dual-task performance in a driving setting. She performed an experiment in which the participants had to drive the LCT and at the same time respond to a visual search task including arrows which pointed in different directions (see Figure 2.3 for examples).

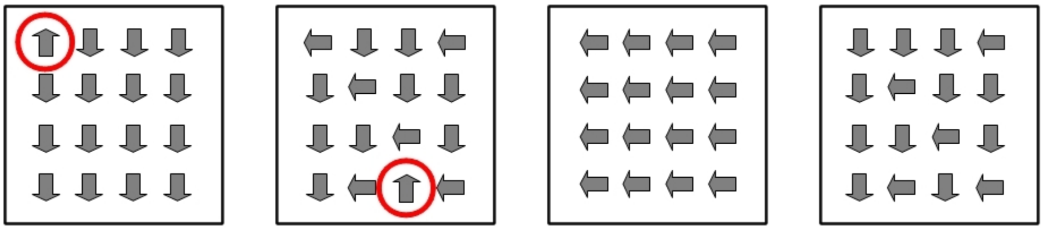


Figure 2.3: Example of the visual search task used in the HASTE project. The encircled symbols represent a target.

It was the participants' task to press a button as fast as possible when one of the arrows pointed upwards. This type of visual search task has, within the European HASTE project, been proven to be a good indicator of the driver's ability to distribute his/her attention between the visual search task and the driving task (Roskam et al., 2002). It can furthermore be argued that arrows are representative for existing IVIS such as navigation systems. However, several concerns with this type of visual task exist. First, potential confounds with using arrows might occur, especially since the LCT requires looking for and responding to arrows as well, causing potential interference due to visual texture segregation and/ or cross-task interference. Second, the arrows in the visual search task might serve as distracting precues, impairing LCT performance. These concerns made us decide to use a different visual search task to make sure, other than the

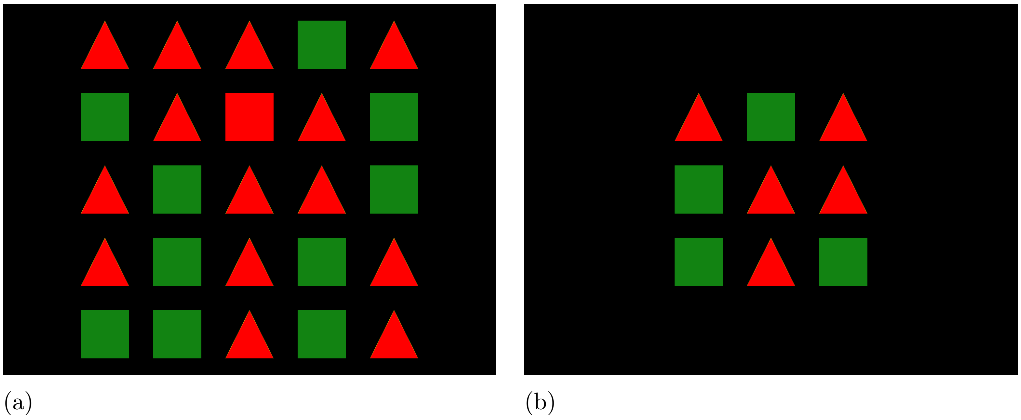

(a) (b)

Figure 2.4: VST examples with (a) and without presence of a target (b).

intended one, no additional interference occured.

We used a conjunction search task first introduced by Treisman and Gelade (1980). Figure 2.4 presents two examples of the visual search display. It consisted of the presentation of a display including distractors (red triangles, red squares) and either one of two types of targets: A green triangle or a red square. It was the task of the participant to decide as fast and as accurate as possible whether a target was present on the display or not. In each case, one of two buttons on the back of the steering wheel, corresponding to the presence or absence of a target item within a set of items, had to be pressed. Task difficulty was manipulated by presenting different sizes of symmetric matrices consisting of 3 x 3 (9) items, 4 x 4 (16) items or 5 x 5 (25) items. Due to use of this task in combination with the driving task, we did not use a fixation cross before presentation of the stimulus display. The reason for this was two-fold: First, we instructed participants to keep their eyes on the driving task and only concentrate on the secondary task when this was needed. Adding a fixation cross might have caused distraction taking the eyes more off the road than necessary. Second, in the dual-task condition, the VST was linked to the progress of the driving task, so a fixation cross could have primed the participants for the possible upcoming of a VST.

Stimuli for the visual search task were generated with E-Prime®. They were presented on a Windows XP 3.00 GHz PC (VST PC: See Figure 2.2 above) driving a 15”

TFT monitor at a resolution of 1024 x 768 pixels (see Figure 2.2). The visual search display was presented at 150 cm in front of the participant and at a visual position of -15 degrees below the horizon of the driving image. At this viewing distance, the display area for the visual search task was a square measuring 3.64° x 3.64° visual angle (°) for the matrix including 3 x 3 items, 5.00° x 5.00° for the 4 x 4 items matrix, and 6.36° x 6.36° for the matrix representing 5 x 5 items. Individual items had a size of 32 x 32 pixels which corresponds to the automotive norms stated in ISO 2575 (ISO 2575, 2004) and ISO 15008 (ISO 15008, 2009) and which represents a visual angle of 0.91° x 0.91°. Vertical and horizontal spacing between each item was 16 pixels (0.46°). VST data were recorded at the end of each trial and contained the trial number, whether a button had been pushed, mean response time, response accuracy and type of search task presented to the participant (target presence, type of target, set size).

### 2.2.3 Dual Task

The dual-task condition consisted of a combination of the LCT and the VST, both as described above. Figure 2.5 presents the timely apparition of the LCT sign and the VST displays in the dual-task condition. One can observe that an additional difficulty was added, as the VSTs were presented either while performing a lane change (1000 ms after visibility of the traffic sign) or while lane keeping (6000 ms after visibility of the traffic sign). Each block of 18 signs in the LCT was thus associated with 36 VST trials. It was emphasized in the written instructions and orally by the experimenter that in the dual-task situation the LCT was always more important than the VST. In case they were not able to perform the LCT and VST at the same time, participants were instructed to prioritize performance on the LCT over performance on the VST.

In the dual-task condition, the timely apparition of the VSTs was assured by the synchronization PC (see Figure 2.2 above): It read out the speed and position information from the LCT and send triggers to the VST PC. By using this method, even in the unlikely event participants slowed down on the LCT, VSTs would always appear at exactly the same moment in time relative to the LCT signs.

To explore eye-movement behavior, electrooculogram (EOG) recordings were made.

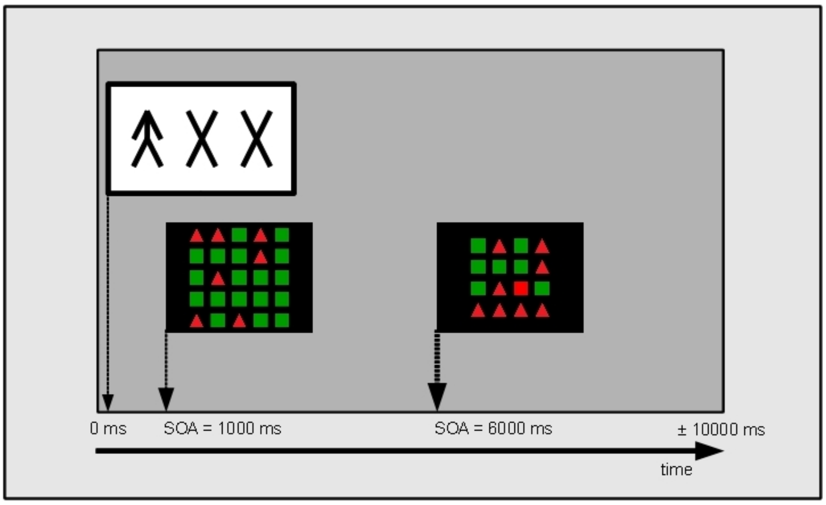


Figure 2.5: Timely apparition of the LCT sign and the VST displays in the dual-task condition.

With EOG recordings no precise areas of interest can be defined but eye glances on and off the road as well as their durations can be measured. For this purpose, 4 Ag/AgCl recording electrodes were placed infra- and supraorbitally in line with the pupil of the dominant eye for vertical (v) EOGs and on the outer canthi of eyes for horizontal (h) EOGs. A fifth recording electrode behind the participant's ear served as reference. The BrainVision Recorder software (Brain Products Inc., Germany), which ran on the *EOG PC* (an IBM compatible PC running Windows XP®) was used to record the EOG signals (see Figure 2.2).

The EOG PC was connected to a BrainVision QuickAmp (Brain Products Inc., Germany). The sampling rate was set to 500 Hz, and the signal was low-pass filtered (200 Hz) and stored for further analysis. As can be seen in figure 2.6, the EOG PC would record EOGs (allowing to count the number of vertical saccades as well as their duration), but serve as a general data server backing up data of the other PCs as well. As such, the type of LCT traffic sign (LC-Sign), its apparition, as well as the optimal track and participant's current position on the track were recorded for the LCT. For the VST, the user response was recorded (VST-Response$_{\text{Onset}}$) as well as the accuracy

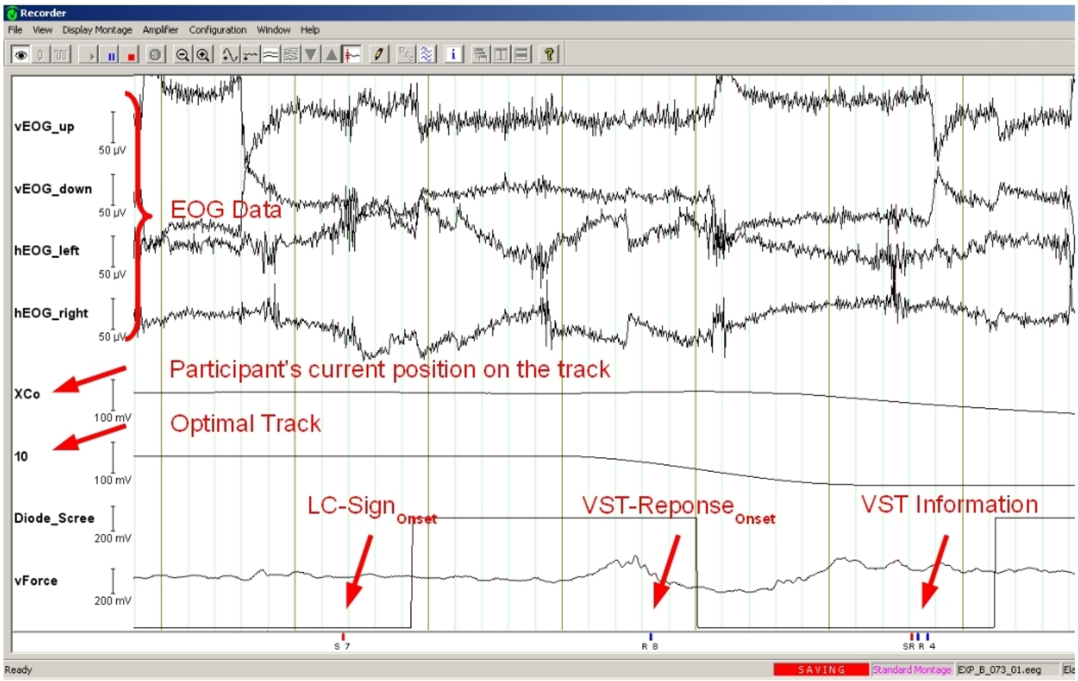


Figure 2.6: Example of the electrooculogram (EOG) PC display. The first 4 channels record eye movements (vertical [v] EOG upper electrode, vEOG lower electrode, horizontal [h] EOG left electrode, hEOG right electrode), the channel XCO and 10 represent information provided by the LCT PC (the participants current position on the track, the optimal track) and the channel Diode_Scree represents a channel with information from the Eprime PC (the diode lighted up when the VST appeared and turned off as soon as a response was detected). Finally there are some events represented, provided by the Eprime PC as well, concerning precise position of the LCT sign (LC-Sign), the moment a button was pressed (VST-Response$_{\text{Onset}}$) and some VST information concerning the type of VST sent (type of target, target present/absent, set size).

of this response and the type of search type (i.e., set size, target type).

### 2.2.4 Standardized Tests

For each experiment, several tests and a demographic questionnaire were administered. In all experiments, a demographic questionnaire was used. Several eye tests were used to measure visual acuity, color blindness (when a VST was part of the experimental setting) and (in experiments with EOG recordings) define the dominant eye. Furthermore, in the experiments with older adults, all participants had to perform two neuropsychological tests to control for aging effects. This section describes each of these tests and questionnaires in detail.

#### Vision Tests

*Ishihara Test for Color Blindness* (Ishihara, 1990). Color perception test used to define whether a participant suffers from red-green color deficiencies. The participant is presented with 15 colored plates, each of which contains a circle of dots appearing randomized in color and size. Participants with normal color vision will discern a number within the pattern of colored dots. For participants with a red-green color vision defect this number will be invisible, or difficult to see. Participants who were incapable of correctly discerning at least 13 of the 15 colored plates were considered red-green color deficient and could not participate in the experiment (as targets for the VST consisted of red and green items).

*Landolt C Visual Acuity Test.* Standardized test for testing vision. This eye test consists of rings (Landolt C) which have a gap either to the left, the right, the bottom or the top. The participant has to decide on which side the gap is located. The size of the C and its gap are gradually reduced until the participant makes a specified rate of errors. The minimum perceivable angle of the gap is taken as measure of the visual acuity.

*Dolman Ocular Dominance Test.* Vision test used to define the participant's ocular dominance. Participants are provided with a card with a small hole in the middle. They are instructed to hold this card with both hands and out-stretched arms in front of them and view at a distant object through the hole with both eyes open. The experimenter then covers one of the two eyes (e.g., left eye) and asks the participant if he/she is still able to see the distant object without moving the position of the card. If the distant object remains visible even with one eye closed, then the open eye is defined as the dominant eye (here: the right eye). The definition of the dominant eye is important for electrode placement, as the electrodes measuring vertical saccades are placed around the dominant eye.

### Trail Making Test (TMT)

The Trail Making Test (TMT) was administered once for each participant at the first experimental session. This neuropsychological test, part of the Halstead-Reitan Test Battery (Reitan & Wolfson, 2003), consists of two tracking tasks. Part A (TMT-A) consists of numbers which have to be connected in an ascending manner. Part B (TMT-B) contains numbers and letters which have to be connected in an ascending manner and with the added task of alternating between the numbers and letters (i.e., 1-A-2-B). Both parts of the test need to be performed as fast as possible and without lifting the pencil from the paper. Part A measures abilities in visuo-spatial scanning and motor sequencing skills, TMT-B allows to measure limitations in executive functioning, psychomotor speed and visual scanning (Lezak, Howieson, & Loring, 2004). Several studies (Ivnik, Malec, Smith, & Petersen, 1996; Rasmusson, Zonderman, Kawas, & Resnick, 1998) have furthermore shown that performance on the TMT declines with increasing age. The critical variable recorded is time to completion. If participants needed more than 78 s on TMT-A or more than 273 s on TMT-B, they were excluded from the experiment.

### Digit Symbol Substitution Test (DSST)

The Digit Symbol Substitution Test (DSST) was administered once for each participant at the first experimental session. This neuropsychological test is part of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1944). Participants are presented with a code table including 9 digits (1-9) each linked to a particular symbol (i.e., 6 - ◯). In a randomized array of digits from 1 to 9 they have to add the corresponding symbol to each digit as fast as possible (see Figure 2.7) from the left to the right. The measured dependant variable is the number of correct symbols within the accorded time of 2 minutes. The DSST gives an indication on the participant's speed of processing including perceptual speed (Gilmore, Royer, Gruhn, & Esson, 2004) known to decline with increasing age (Salthouse, 1992).

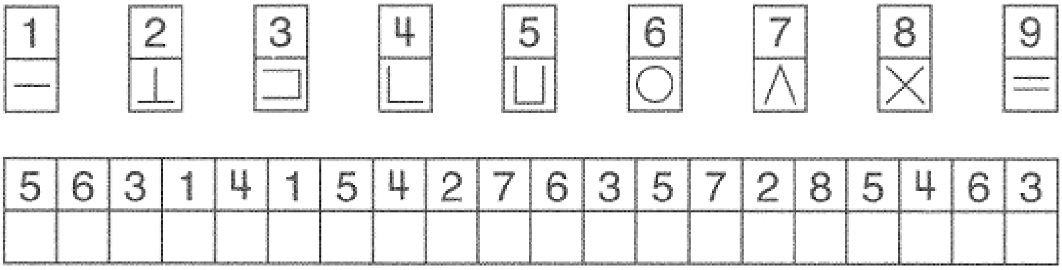


Figure 2.7: Code table used in the Digit Symbol Substitution Test with digit-pairs (above) and an array of random digits from 1 to 9 (below) with space below where the participant has to add the correct symbol according to the code table.

### Rating Scale of Mental Effort (RSME)

For a subjective measure of workload after each track, we employed the Rating Scale of Mental Effort (RMSE; Zijlstra & Van Doorn, 1985). This simple one-dimensional ratio scale allows participants to rate their invested mental effort into a task on a scale from 0 (absolutely no effort) to 150 (extreme effort).

### NASA Task Load Index (NASA-TLX)

The NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) was filled out at the end of a series of experimental blocks of the same type (e.g., after 5 VST experimental blocks). This multidimensional questionnaire is intended to assess overall mental workload experienced while performing a task according to six dimensions: Mental demand, physical demand, temporal demand, performance, effort, and frustration. *Mental* *demand* assesses how much mental and perceptual activity (e.g., thinking, calculating, looking, searching, deciding) was required for performing the task. *Physical demand* measures the amount of physical activity that was needed (e.g., pushing, pulling, controlling, activating) to perform the task. *Temporal demand* assesses whether the participant felt under time pressure while performing the task. *Effort* rates how hard a participant estimates he had to work (mentally and physically) to accomplish his/her performance. *Performance* describes the degree of success or satisfaction felt upon the performance or completion of a given task. A high score on this scale means that the participant rates his own performance as poor, a low score as good. And finally *frustration* assesses to which extent a participant felt insecure, discouraged, stressed and irritated while performing the task.

## 2.3 Design and Procedure

Table 2.1 presents the different experiments and their conditions, the number of experimental trials per condition as well as the number of sessions per experiment. Depending on the experiment, participants were either tested in a single session or in a multi-session setting. Each experimental session consisted of the single-LCT, single-VST and the dual-task condition.

Table 2.1: Number of experimental blocks and sessions as a function of experiment (1-4) and task (single-LCT, single-VST, Dual Task).

| Experiment | single-LCT | single-VST | Dual Task | single-LCT | single-VST | Sessions |
|---|---|---|---|---|---|---|
| 1 | 5 | 5 | 10 | 5 | 5 | 1 |
| 2 | 5 | 5 | 10 | 5 | 5 | 1 |
| 3 | 5 | 5 | 10 | 5 | 5 | 4 |
| 4 | 5 | 5 | 2 x 10 | – | – | 4 |

In all experiments, participants first read and signed a consent form and filled out the demographic questionnaire before performing the TMT-A, TMT-B and the DSST. For experiments that included EOG recordings, the 5 electrodes were then placed on the participants' face (see above for details).

A customized demographic questionnaire was submitted once to each participant and assessed their age, sex, dexterity, mean number of km driven per year and the number of years they were in possession of a valid driver license. The questionnaire furthermore collected some general information concerning the use of driver assistance systems, such as navigation systems, cruise controls, adaptive cruise controls and lane-departure warning systems.

They were then seated in a quiet, darkened room in which they received written instructions describing the LCT. Participants were informed to change lanes as fast and as accurately as possible as soon as the initially blank sign became visible (at 40 m before the actual position of the sign), but not before, and to keep a constant driving speed of 60 km/h throughout the drive (which consists of pushing the accelerator pedal to its maximum, as the simulation then automatically kept a constant speed of 60 km/h). Ten tracks with different sign orders are available within the LCT. A "track" was considered an experimental "block" and therefore those terms will be used interchangeably. It was furthermore emphasized that good performance included lane-change speed and accuracy, as well as lane keeping in the middle of the lane and when driving straight. Each participant drove then at least 1 and up to 3 practice blocks. The participants then had to drive 5 LCT experimental blocks. After each block of 18 LCT trials, participants had to rate their mental effort on the RSME. At the end of the LCT experimental block, each participant had to fill out the NASA-TLX and rate his/her general mean subjective workload for the last 5 tracks.

Participants then received written instructions describing the visual search task, in which it was emphasized to decide as quickly and as accurately as possible whether a target was present or absent. Each experimental block consisted of 36 VST trials: 2 targets (present/absent) x 2 target types (red square/green triangle) x 3 set sizes (9, 16, 25 items) x 3 repetitions. The position of the "target present button" was counter-balanced over participants. All trials were presented in a pseudo-random sequence. Target duration was 2500 ms followed by an inter-trial interval of 2500 ms. In the practice block, when participants pushed the wrong button, they received immediate vocal feedback saying "falscher Hand" ("wrong button"). In case of a miss, a vocal feedback followed saying "zu spät" ("too late"). When the correct answer was given (the correct button pushed) the visual search task disappeared upon response. After each block, participants were presented with feedback concerning their mean reaction time (ms) and accuracy (%) on the VST. Each participant performed then at least 1 practice block at the end of which they received feedback on their performance. Generally 1 practice block was needed for all age groups to understand the task at hand (and reach an accuracy level of 70%). The participants then had to perform 5 experimental visual-search task blocks. After each block, participants had to rate their mental effort on the RSME. At the end of the 5 VST experimental blocks, each participant had to fill out the NASA-TLX and rate their general mean subjective workload for the last 5 VST blocks.

They then received written instructions describing the dual-task condition. Not only was it emphasized that performance on both the LCT as well as the VST should be as fast and as accurate as possible, but special emphasis was put on the higher priority of the driving task. Each participant performed then at least 1 practice block at the end of which they received the same feedback as in single-task LCT and VST conditions mentioned above. Generally, younger participants needed 1 practice block to understand the task at hand, whereas older adults sometimes needed up to 3 practice blocks. The participants then had to perform 10 experimental dual-task blocks. After each block, participants had to rate their mental effort on the RSME. At the end of the 10 dual-task experimental blocks, each participant had to fill out the NASA-TLX and rate their general mean subjective workload for the last 10 dual-task blocks.

Finally participants received once more written instructions describing the LCT, before performing 5 experimental single-task LCT blocks, with a rating of their mental effort on the RSME after each block and the NASA-TLX at the end of the 5 experimental blocks. Finally they would receive written instructions describing the VST before performing 5 experimental single-task VST blocks as well. They would again rate their subjective mental effort on the RSME after each block and rate their general subjectively rated effort over the last 5 VST blocks on the NASA-TLX.

The participants were allowed to rest in between practice or experimental blocks. All experimental sessions took approximately 3 hours.

## 2.4 Data Analysis

This section gives an overview of all dependent variables (DV) used throughout the experiments and for each experimental task. Note that EOG measures as well as mean RSME and mean NASA-TLX scores are not separate tasks, but as they play a role in data analysis and interpretation, small subsections have been dedicated to those measures within this data-analysis section.

Probability level for statistical significance of all analyses was set to .05 and sphericity violations were corrected for with the Greenhouse-Geisser $\varepsilon$. However, for readability's sake, uncorrected degrees of freedom are provided. The reported effect size $\eta_p^2$ (partial eta-squared) is defined as the proportion of the effect plus error variance that is attributable to the effect.

### 2.4.1 LCT

Following the work of Huemer and Vollrath (2010) and Maciej and Vollrath(2009) we separated each LCT trial in a Lane-Change (LC) and Lane-Keeping (LK) segment. The LC segment consisted of the actual lane change period following a sign indication, whereas the LK segment consisted of the period just before a new sign appeared and in which the participant only had to drive straight on. The classical LCT analysis software does not distinguish between LC and LK segments, but averages driving performance measures over those segments. Although this method has been proven valuable especially when comparing between conditions (for examples see Harbluk et al., 2007, 2009; Mitsopoulos-Rubens, Trotter, & Lenné, 2010; Young, Lenné, & Williamson, 2011), in some cases effects might be under- over-estimated or even disappear completely. Analyzing driving performance measures of the LCT based on LC and LK segments presents the advantage of being more precise.

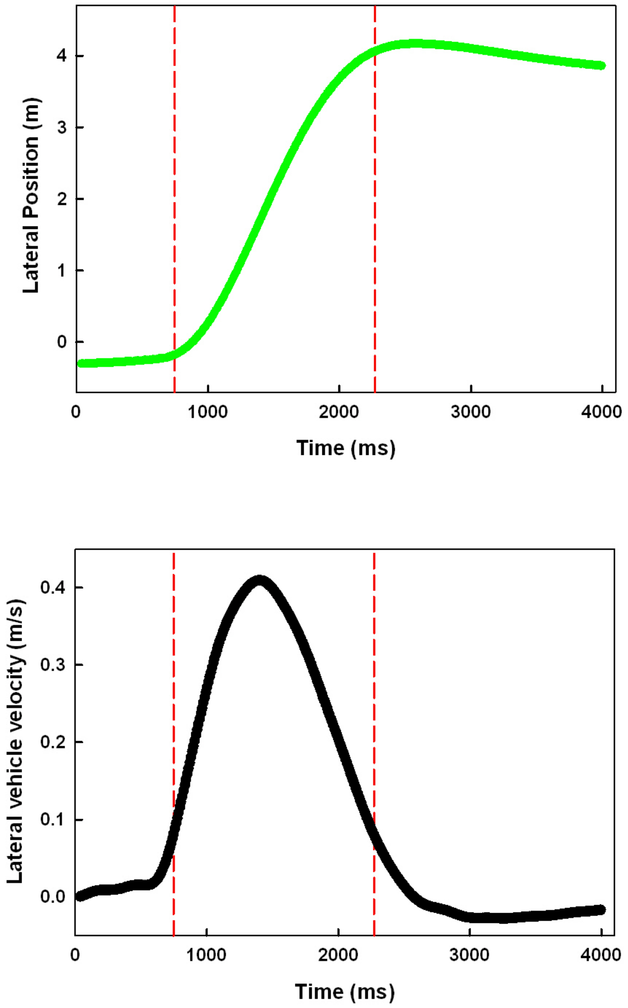


Figure 2.8: Representation of a lane-change maneuver (upper figure) and the resulting numerical derivation (lower figure) based on which the beginning ($\text{LC}_{\text{Onset}}$) and the end ($\text{LC}_{\text{Offset}}$) of a lane-change movement are calculated. The x-axes of both figures represent the time on the trial in milliseconds (ms). The y-axis of the upper figure represents the lateral position in meters (m). Zero (0) represents the middle of the middle lane. The y-axis of the lower figure represents the lateral vehicle velocity in meters per second (m/s). The red dashed lines in both figures represent the moments in time at which lateral vehicle velocity reached 20% before and after the peak lateral vehicle velocity.

To define the exact duration of a LC segment, we used lateral vehicle velocity to define the beginning ($\text{LC}_{\text{Onset}}$) and the end ($\text{LC}_{\text{Offset}}$) of the lane-change maneuver based on the actual movement of the participant. Figure 2.8 presents an example of a lane-change maneuver (upper figure) and the resulting numerical derivation (lower figure). In order to calculate $\text{LC}_{\text{Onset}}$ and $\text{LC}_{\text{Offset}}$, first a baseline correction was applied to account for potential drifts in lateral position prior to a lane change. This baseline was established by taking the average of the 10 last samples prior to $\text{LC-Sign}_{\text{Onset}}$. This baseline was then subtracted from the subsequent lateral positions (until the next $\text{LC-Sign}_{\text{Onset}}$) and lateral vehicle velocity was then obtained by numerical derivation of the lateral vehicle position over time. $\text{LC}_{\text{Onset}}$ and $\text{LC}_{\text{Offset}}$ were defined as the first moments in time at which lateral vehicle velocity reached 20% of peak lateral velocity before and after this peak, respectively. Figure 2.9 presents different points in time within one trial and how they combine to calculate not only LC and LK segments, but independent temporal variables for the LCT, the VST and the dual-task condition as well. As mentioned in the introduction of this subsection, LC segments consist of the total time from visibility of a traffic sign ($\text{LC-Sign}_{\text{Onset}}$; set to 40 m before the actual passing of the sign) until the end of the lane-change movement ($\text{LC}_{\text{Offset}}$).

Duration of LK segments consist of the total trial time (from the $\text{LC-Sign}_{\text{Onset}}$ to the next $\text{LC-Sign}_{\text{Onset}}$) minus the duration of the LC segment minus a buffer to avoid the inclusion of later portions of the LC in LK segments, as shown in Figure 2.9.

LK Segment Duration = Trial Duration - LC Segment Duration - Buffer Duration

The separation in LC and LK segments not only has the advantage of analyzing LCT data in a more detailed manner, but allows as well the creation of some additional dependant variables such as Reaction Time until Lane Change (RT-LC) and Movement Time Lane Change (MT-LC). In what follows we will have a closer look at some new as well as classical LCT performance measures.

*Reaction Time until Lane Change (RT-LC).* RT-LC is the response time of a participant to start the actual lane change after an LCT traffic sign gets visible. $\text{LC-Sign}_{\text{Onset}}$ marks the beginning and $\text{LC}_{\text{Onset}}$ marks the end of this period. This measure examines how long it takes a participant to start the actual lane-change maneuver. One can expect that factors such as interference from secondary tasks or age effects influence this

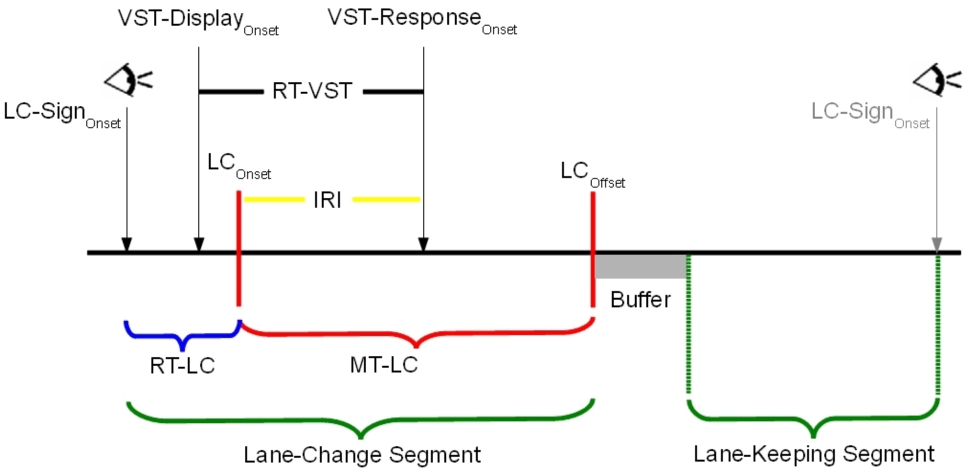


Figure 2.9: Representation of different points in time within one trial and how they combine to account for relative lane-change and lane-keeping segments, as well as independent temporal variables for the LCT, the VST and the dual-task condition (see text for details).

measure. RT-LC is a relatively little used performance measure for the LCT, but it has the advantage to serve as an event-detection measure providing a reaction time to a precise event (Harbluk et al., 2007; Young et al., 2011) and allows explaining certain confounding effects observed in MDEV: If participants take more than 600 ms to initiate a lane change, this has a negative effect on MDEV, as the optimal track changes of lane 600 ms after a LCT-sign becomes visible. It has been used before by Benedetto et al. (2011) under the name *Lane Change Delay*, by Huemer and Vollrath (2010) under the name *Reaction Time (Lane Change)* and by several other authors (Bruyas et al., 2008; Harbluk et al., 2007; Young et al., 2011) under the name *Lane Change Initiation*. As RT-LC is directly related to the lane-change maneuver, it can be calculated for LC segments only.

*Movement Time Lane Change (MT-LC).* MT-LC represents the period of time in which a participant is actually changing of lane. In Figure 2.9 it can be seen that MT-LC starts at $LC_{Onset}$ and ends at $LC_{Offset}$. It has the advantage of representing an event measure which is independent of the optimal track defined by the LCT. As MT-LC is directly related to the lane-change maneuver, it can be calculated for LC segments only.

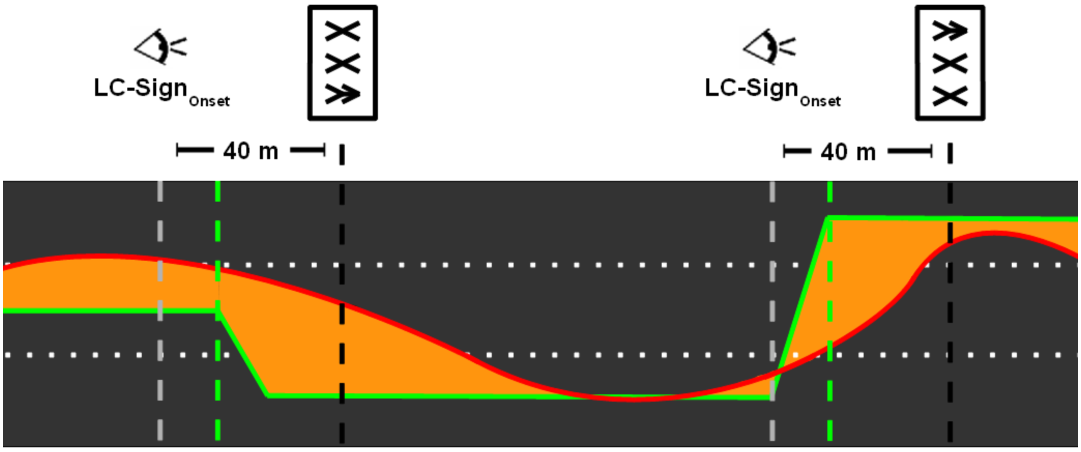


Figure 2.10: Illustration on how lateral deviation measures are obtained: The green line represents the reference trajectory. The red line represents the driving behavior of the participant and his deviation from the reference trajectory (orange area). LCT signs (black vertical dashed lines) are indicated at their actual position. The eyes as well as the grey vertical dashed lines represent visibility of the LCT signs (LC-Sign$_{\text{Onset}}$ at 40 m before the actual position of the LCT sign. The green vertical dashed lines illustrate the moment at which the reference trajectory starts the lane change (at 30 m before the actual position of the LCT sign).

*Mean Lateral Deviation (MDEV).* This measure provides information about lateral control on the LCT that is obtained by comparing the actual driving trajectory of the participant to a reference trajectory (Mattes & Hallén, 2009). The normative model is the same for each participant. It assumes that participants take 600 ms to respond to the apparition of a traffic sign (which corresponds to 10 m on the track) before making a lane change within 10 meters (independant of a single or a double lane change). Figure 2.10 illustrates how lateral deviation measures are obtained for the calculation of MDEV. The green line represents the reference trajectory. The red line represents the driving behavior of the participant and his deviation from the reference trajectory (orange area). LCT signs are indicated at their actual position. The eyes represent visibility of the LCT signs (LC-Sign$_{\text{Onset}}$ at 40 m before the actual position of the LCT sign). The moment at which the reference trajectory starts the lane change (30 meters before the actual position of the LCT sign) is indicated by the green vertical dashed lines. MDEV consists of the mean of the absolute values of the lateral deviations over time. As MDEV is a combined measure of object and event detection (Harbluk et al., 2009), one needs to keep in mind the potential confound with the participant's reaction time to a lane-change sign when analyzing MDEV: The longer it takes to react and start the lane-change maneuver, the higher MDEV will be. This effect might be especially notable when prompt reaction to a lane-change sign is impaired by presentation of an interfering secondary task just after apparition of this sign. MDEV is thus especially of interest for comparing conditions (e.g., single vs. dual-task conditions) as absolute values provide limited information (Mitsopoulos-Rubens et al., 2010). MDEV can be calculated for both LC as well as LK segments.

*Standard Deviation of Lateral Deviation (SDDEV).* This measure provides information about the lateral-deviation variability when lane changing or lane keeping. SDDEV is computed by taking the standard deviation of the absolute value of the lateral deviations over time. It has been shown to be a powerful indicator of driving performance especially in lane-keeping segments (Berthon-Donk et al., 2011). SDDEV can be calculated for both LC as well as LK segments.

*Standard Deviation of the Steering Wheel angle (SDSW).* SDSW provides information about the variability in steering wheel variation. It reflects the number of corrective steering movements which are made to perform a lane-change or lane-keeping maneuver. SDSW often shows a behavior which is opposed to performance measures such as MDEV and/or SDDEV. Berthon-Donk et al. (2011) found that, although MDEV improved over blocks, this was paired with an increase in steering-wheel variability. The authors argued that these findings were consistent with findings from Latash, Scholz, and Schöner (2002) that variability also reflects the ability to compensate for unintended deviations. SDSW can be calculated for both LC and LK segments.

Before analysis of LCT-data, several exclusion criteria were applied to assure proper data. A complete trial (including the LC and LK segment) was discarded when any of the following hierarchical criteria were met. First, algorithm errors were excluded (*algorithm filter*). An algorithm error occured, when the software failed to detect the moments in time at which lateral vehicle velocity reached 20% of peak lateral velocity before and/or after the peak (this occurs for example when a participant fails to change of lane or when his lane-change maneuver occurs with an extreme delay). Second, when RT-LC was smaller than 400 ms or bigger than 2000 ms a trial was discarded (*RT-LC filter*) as well. This filter was followed by the *MT-LC filter* according to which trials were discarded when MT-LC values were smaller than 1000 ms or bigger than 6000 ms. Both RT-LC and MT-LC exclusion criteria were based on critical visual inspection of pilot data from a representative data set including younger and older adults. The *LK filter* filtered out trials in which the LK time was smaller than 1000 ms. The *LC-Error filter* filtered out trials in which participants had made an incorrect lane change or had deviated into another lane. A lane change was considered incorrect if, at the next visible sign, participants were not in the lane they were supposed to be based on the previous sign. And finally the *LC-Miss filter* filtered out data in which the participant had ignored the LCT sign. A lane-change sign was considered missed, when participants were in the same lane as they were at the end of the previous lane-change maneuver.

### 2.4.2 VST

The participants' performance on the VST was measured by the RT on the VST (RT-VST) and the number of errors produced. As can be seen in Figure 2.9, in the dual-task condition the VST is presented at a fix interval after apparition of the traffic sign (VST-Display$_{\text{Onset}}$). The VST remains visible until the participant pushes a button, or after a time-out of 2500 ms (VST-Response$_{\text{Onset}}$). The time a VST display remains visible represents the RT-VST. Errors were divided into *incorrect responses* (the participant provided an inaccurate response) and *misses* (the participant failed to push a button within 2500 ms).

For both errors as well as RT-VST only trials in which a target was present were used. Furthermore, data for which no correct response was given (wrong button presses/misses) were excluded for analysis of RT-VST.

### 2.4.3 Dual Task

The participants' performance in the dual-task condition was analyzed using all dependant variables from the LCT and VST (filtered as described above) as well a as a measure indicating potential interference from the secondary task on the primary driving task, EOG measures and subjective measures.

*Inter-Response Interval (IRI).* This measure provides information on the interaction between the LCT and the VST in dual-task conditions. Figure 2.9 shows that IRI represents the time it takes participants to respond to a VST (VST-Response$_{\mathrm{Onset}}$) starting from the moment they initiated their lane-change maneuver (LC$_{\mathrm{Onset}}$). An increase in IRI represents an increase in interference from the VST on the LCT. Note that for consistency reasons with the VST, we only analyzed IRI trials in which a target had been present in the VST and if a correct response on the VST was given. It should furthermore be noted that IRI can be calculated for LC segments only.

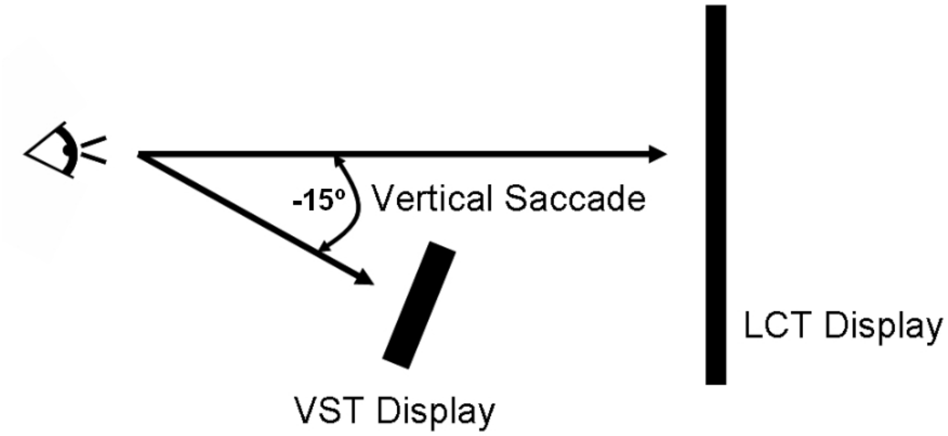


Figure 2.11: Vertical saccade between the LCT display and the VST display.

EOG recordings allowed the extraction of vertical saccades by analyzing EOG traces (see Figure 2.11). First a differential vertical channel was calculated by combining the upper and lower vertical (v)EOG channels. This channel was then segmentated to filter for sections in which a target had been present, a button had been pressed by the participant and in which a correct answer was given. After a baseline correction, markers for onset and offset of an eye movement were set when a threshold of 100 microVolt was reached. Once the vertical saccades extracted from the EOG channel, their frequency as well as their duration could be quantified.

*Vertical Saccade Frequency (VSF).* This measure provides information on the number of times participants take their eyes off the road to look at the VST display within a given period. A vertical saccade is defined as a downwards glance towards the VST display, where each vertical saccade is separated by at least one vertical saccade back to the road.

*Fixation Duration (MFD).* This measure provides information on the period of time participants take their eyes off the road to look at the VST display. Individual fixation duration is defined as the vertical movement downwards from the road towards the visual search display until a vertical movement from the visual search display upwards to the road again.

*Total Glance Duration (TGD).* Measure which provides the total glance time towards the visual search display, obtained by multiplying the number of vertical saccades with the fixation duration.

VSF and MGD were analyzed for lane-change and lane-keeping segments and, to be consistent with the VST, only for segments in which a target was present and a correct response was given.

### 2.4.4 Subjective Measures

These measures provided insight into the perceived subjective load based on rating scales (see the subsection RMSE, Chapter 2.2.4 for a more detailed description on those measures). Both mean RSME and mean NASA-TLX were analyzed for all task conditions (LCT, VST, dual-task). NASA-TLX was analyzed according to the 6 individual subscales present in the rating scale. Except for incomplete data, no data was excluded for analysis.

# 3 Experiment 1: Methodological Study

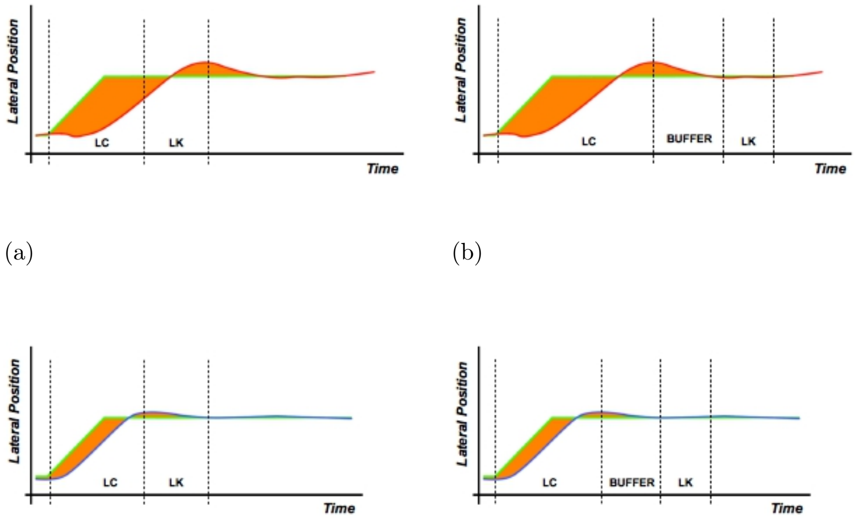
## 3.1 Introduction

In recent years, the Lane Change Test (LCT; Mattes, 2003) has been established as a reliable tool for measuring in-vehicle task demands caused by navigation systems (Harbluk et al., 2007, 2009; Maciej & Vollrath, 2009), music players (Mitsopoulos-Rubens et al., 2010) or cognitive tasks serving as surrogate in-vehicle information systems (Benedetto et al., 2011; Engström, Johansson, & Östlund, 2005; Wilschut et al., 2008; Young et al., 2011). The test consists of a simplified driving environment in which lane changes need to be made in response to traffic signs. Due to its nature, the LCT corresponds to a tracking task and performance is measured by deviation from an optimal track. The amount of distraction due to a secondary task is measured by an alteration in lane-change performance while dealing with a secondary task as compared to lane-change performance without secondary task (Mattes & Hallén, 2009).

Although the LCT is a useful tool for measuring driver distraction, the issue of how to analyze data from the LCT remains partially to be clarified. As mentioned in the Data Analysis section of the General Method (Chapter 2.4). The classical LCT analysis software does not distinguish between LC and LK segments, but averages driving performance measures over those segments. However, considering the schematic illustration of a lane-change maneuver in figure 2.10 (Chapter 2.4), it may intuitively be expected that some parts of the LCT to be more difficult (i.e. LC segments) than others (i.e. LK segments). It thus seems useful to analyze LCT performance measures according to LC and LK segments. One question that needs to be addressed is how to define those segments. Maciej and Vollrath (2009) as well as Huemer and Vollrath (2010) used absolute windows as a method to differentiate between LC and LK segments. In both studies, a LC segment was defined to begin at 30 m before the actual position of the sign and lasted 80 m (until 50 m after the actual position of the sign). LK segments were defined to start 50 m before visibility of the traffic sign and lasted 50 m. Dividing the track in LC and LK segments allowed to analyze driving measures more in detail and explain certain effects which otherwise would have stayed unnoticed (e.g., the influence of driving task complexity). Berthon-Donk, Grosjean and Rinkenauer (2011) used absolute windows as well to differentiate LC from LK segments in a study that assessed the effect of feedback on LCT performance measures. Their LC segment was defined to begin at 40 m before the actual sign position and lasted 96 m. LK segments were defined to start 40 m before visibility of the LC-Sign and lasted 40 m. By using those absolute windows, they could successfully show that separating lane-change maneuvers in different driving segments, yielded driving performance outcomes that better represented differences in lane change and lane keeping.

One problem of using absolute windows to calculate driving performance measures is the definition of segment length. As can be seen by the studies mentioned above, different authors use different segment lengths leading to potential differences in calculated driving measures and making study outcomes hard to compare. Another problem of using absolute windows to calculate driving performance measures is that some driving performance measures can easily be over- or underestimated, as the length of absolute windows might not be suitable to reflect actual driving behavior.

In what follows, different examples of lane-change maneuvers either separated into LC and LK segments by using absolute windows, or by using relative windows are considered. Figures 3.1a and 3.1b present an example for a slow lane-change maneuver (as could be the case for an older person). In Figure 3.1a, absolute windows are used to define LC and LK segments: The LC segment has a predefined length with a fix start and ending point. The LK segment starts immediately at the end of the LC segment and has a predefined length and a fix start and ending point as well. It can be observed that the absolute LC segment is not adapted to the actual length of the lane-change maneuver: Parts of the LC maneuver are actually in the LK segment. For this reason, lane-change accuracy data (e.g., MDEV-values) for the LC segment might be artificially underestimated and lane-change accuracy data for the LK segment might be overestimated as it includes LC data. Figure 3.1b represents this exact same lane-change maneuver, but LC and LK segments are defined using a relative calculation method, that is, based on the



(a) (b)

(c) (d)

Figure 3.1: Examples of different types of lane-change maneuvers divided into Lane Change (LC) and Lane Keeping (LK) segments either by using absolute windows (a and c) or by using relative windows (b and d). The x-axis represents time, the y-axis represents the lateral position of the vehicle. The green line represents the reference trajectory. The red line represents the driving behavior of a slow participant, the blue line represents the driving behavior of a fast participant. The orange area represents the deviation from the reference trajectory. The vertical dashed lines illustrate the beginning and the end of LC and LK segments. Note that these figures serve as examples and are not to scale.

actual lane-change behavior of the participant. As can be seen, the LC segment starts at the same point as the absolute window, but LC segment-duration differs and is more in accordance with the actual lane-change length of the participant. The LK segment, due to the integrated buffer (See Chapter 2.4), does not take into account data points from the LC segment. Lane-change accuracy data calculated with the relative windows will in this example be higher than lane-change accuracy calculated with absolute windows and lane-keeping accuracy data calculated will in this example be lower than lane-keeping accuracy data calculated with absolute windows due to inaccurate segment lengths (for LC) and segment positions (for LK).

Figure 3.1c and Figure 3.1d present an example for a fast lane-change maneuver (e.g., for a younger person). In Figure 3.1c, absolute windows are used to define LC and LK segments. It can be observed, that data points belonging to the LK segment (i.e., driving straight), are included in the LC segment. Resulting lane-change accuracy data (e.g., MDEV) in the LC segment, might for this reason be underestimated when using an absolute window, as data points from the LK attenuate performance measures. Figure 3.1d presents this same lane-change maneuver with relative windows for the definition of LC and LK segments. The LC segment starts at the same point as the absolute window, but has a length which is adapted to the lane-change maneuver of the participant. For this reason, the LC segment better corresponds to the actual lane-change maneuver of the participant. Lane-change accuracy data calculated with the relative window for LC will therefore be higher as compared to performance outcomes calculated with an LC segment based on an absolute window. In this example, the use of absolute versus relative windows for the calculation of the LK segment has no effect on data in the LK segment.

The problem described with the example above, might be especially of relevance when different age groups are tested. There is now a lot of evidence that aging influences the motor-sensory system that affects movement abilities as evidenced by a slowing in tracking tasks (Jagacinski, Liao, & Fayyad, 1995), arm movement tasks (Pohl, Winstein, & Fisher, 1996) as well as reaction times (Alm & Nilsson, 1995; Rubichi et al., 1999; Shanmugaratnam et al., 2010). The LCT not only requires precise and rapid tracking involving rapid arm movements, but also rapid reaction times as soon as a traffic sign appears. Due to general slowing of the sensory-motor system for older adults, an increase in reaction time can thus be expected, resulting in lane-change movements starting at a later time as compared to younger adults. Furthermore, due to a general slowing in motor functions, lane changes might take longer for older adults as compared to younger adults. Defining LC and LK segments based on absolute values might be inaccurate for this particular group of participants.

Recent movement studies suggest (e.g., Sülzenbrück & Heuer, 2011), that adaptive reference tracks (i.e., reference tracks that are adapted to the individual movement pattern of the participant) are more precise for the calculation of movement trajectories and the resulting dependent variables. It can thus be expected that when lane-change and lane-keeping segments are defined by relative boundaries, based on individual kinematics while lane changing, the performance metrics for the LCT will be more precise and will allow for better adaptation to different populations. More particularly, the use of relative windows for the definition of LC and LK segments more accurately reflects driving behavior in each of the driving segments. Performance measures in LC segments will not be attenuated by the inclusion of data points from LK segments, and performance measures in LK segments will not be overrated due to the inclusion of data points from LC segments. To our knowledge, no study has investigated whether relative calculation methods can be used for the definition of LC and LK segments in the LCT and what their effect is on driving performance measures.

### 3.1.1 Research Questions and Hypotheses

The goals of this study were threefold. First, the question whether the use of a relative calculation method can be used for defining LC and LK segments was examined. As the lane-change maneuver is a movement trajectory, it was expected that a relative calculation method can be useful to define LC and LK segments based on individual lane-change behavior.

Second, the effect of using absolute versus relative LC and LK windows on driving-performance measures in the LCT was examined. As performance metrics with relative LC and LK windows will be adapted to actual lane-change behavior, it is expected that they will yield more representative performance measures.

Third, as this thesis is focused on the comparison between younger and older adults, it was examined more into detail whether the use of relative windows has an effect on LCT performance measure outcomes when comparing both populations (in comparison to absolute windows). Based on recent aging studies in driving contexts (Fofanova &

Vollrath, 2011), it can be assumed that lane-change behavior of older adults differs from that of younger adults. If the use of relative LC and LK windows more accurately reflects driving behavior (i.e., performance results are neither attenuated or enlarged by the inclusion of irrelevant data points, nor reduced or expanded by the exclusion of relevant data points), it is expected that age differences will be reflected more accurately in the findings calculated with the relative windows as compared to the absolute windows.

## 3.2 Method

### 3.2.1 Participants

A total of 23 individuals participated in the experiment. The data of 1 younger participant had to be discarded due to technical problems. Eleven younger participants (mean age = 24.0 years, age range = 21-27 years) and 11 older participants (mean age = 71.2 years, age range = 64-80 years) remained for analysis. Younger drivers reported driving 10454 km per year and were in possession of a driver license for a mean period of 6.0 years. Older participants reported driving 9545 km per year and their mean period of possession of a driver license was 46.1 years.

### 3.2.2 Procedure

The single-session experimental design was applied as described in the General Method section (2.3).

### 3.2.3 Data Analysis

Because the main focus of this study was to explore how different calculation methods for LC and LK segments affect driving performance measures, data were analyzed on the basis of those two calculation methods. The first method of defining LC and LK segments is relative, that is, based on the actual lane-change behavior of the participant. Note that the term "adaptive" is not used, as it has already been used in LCT literature for another concept (see for example Minin, Benedetto, Pedrotti, Re, & Tesauri, 2012; Tattegrain, Bruyas, & Karmann, 2009). The length of the LC segment is calculated based on a numerical derivation which defines the beginning and the end of the lane-change movement. As mentioned in Data Analysis section of the General Method 2.4, LC segments consist of the total time from visibility of a traffic sign (set to 40 m before the actual passing of the sign) until the end of the lane-change movement. This method will throughout the results section be referred to as "*relative*". The second method, called "*absolute*" defined LC and LK segments by the use of absolute windows. The absolute LC segments were defined to start with visibility of the traffic sign (40 m before the actual position of the sign) and to end 56 m after the actual position of the traffic sign, for a total distance of 96 m. Those cut-off criteria were based on critical visual inspection of pilot data from a representative data set including younger and older adults. LK segments were defined to start 80 m before the actual position of the following sign and to end 40 m before that same sign, for a total of 40 m, which corresponds to LK-segment lengths used by Huemer and colleagues (Huemer & Vollrath, 2010).

LCT trials analyzed with the relative calculation window were excluded following the different filters described in the Data Analysis section 2.4 of the General Method. For the remaining older adults ($N = 11$), the mean percentages of excluded LCT trials were 1.34% (algorithm-error filter), 8.47% (RT-LC filter), 1.06% (MT-LC filter), 0.37% (LK filter), 5.37% (wrong-LC filter) and 0.05% (missed-LC filter), for a total of 16.67% discarded trials in the dual-task condition. The total percentage of discarded trials in the single-LCT condition added up to 3.68% for the older adults. For the remaining younger adults ($N = 11$) 1.31% of all dual-task LCT trials were discarded and 0.91% of the single-LCT trials. LCT trials analyzed with the absolute window were excluded following the wrong-LC filter and missed-LC filter only. For older adults, they represented 8.66% and 0.09% respectively in the dual-task condition and 0.96% and 0.00% respectively in the single-LCT condition. For younger adults, they represented 0.71% and 0.00% respectively in the dual-task condition and 0.30% and 0.00% respectively in the single-LCT condition. Although visual search task (VST) trials were included in the experiment, to compare between single- and dual-task conditions, VST data were not analzyed.

Separate 4-way mixed-factors ANOVAs were conducted on the driving measures MDEV, SDDEV and SDSW. For a more detailed description of these ANOVAs see Appendix A.

## 3.3 Results

In this results section, due to a large amount of data, only results of relevance for the research questions are presented. These results include mainly the factor *window*, to examine the effect of using different calculation windows on performance measure outcomes and the factor *age*, to examine the effect of age in general. Significant interactions including (at least) the factors window and/or age are presented as well, as these can provide us with useful information on how those factors interact among them or with other factors. An overview of the complete data analysis, can be found in Appendix A. Due to the amount of data and to keep the results section concise, figures were only included for significant or close-to-significant ($p$ =.08) main effects or interactions. Note that close to significant interactions will only be described, without providing any exploratory follow-up tests.

*MDEV.* The ANOVA revealed a significant main effect of window ($F[1, 21] = 220.18$, $\eta_p^2 = .91$, $p < .001$): MDEVs calculated with relative windows were higher (M = .88 m) than those calculated with absolute windows (M = .71 m). There was furthermore a significant main effect of age ($F[1, 21] = 36.71$, $\eta_p^2 = .63$, $p < .001$): Older adults yielded significantly higher MDEV values (M = .98 m) than younger adults (M = .61 m).

Figure 3.2 presents mean MDEV as a function of task type, window and age group. There was a significant 2-way interaction between task type and age ($F[1, 21] = 5.05$, $\eta_p^2 = .19$, $p < .05$) as well as task type and window ($F[1, 21] = 12.91$, $\eta_p^2 = .38$, $p < .01$). These interactions were however modulated by a significant 3-way (age x task type x window) interaction ($F[1, 21] = 5.38$, $\eta_p^2 = .20$, $p < .05$). To follow up on this interaction we performed two separate 2-way (task type x window) ANOVAs for each age group respectively. There was no significant interaction between those two factors for the younger adults ($p > .10$), but the interaction reached significance for the older adults ($F[1, 11] = 10.76$, $\eta_p^2 = .49$, $p < .01$). A follow-up on the 2-way interaction

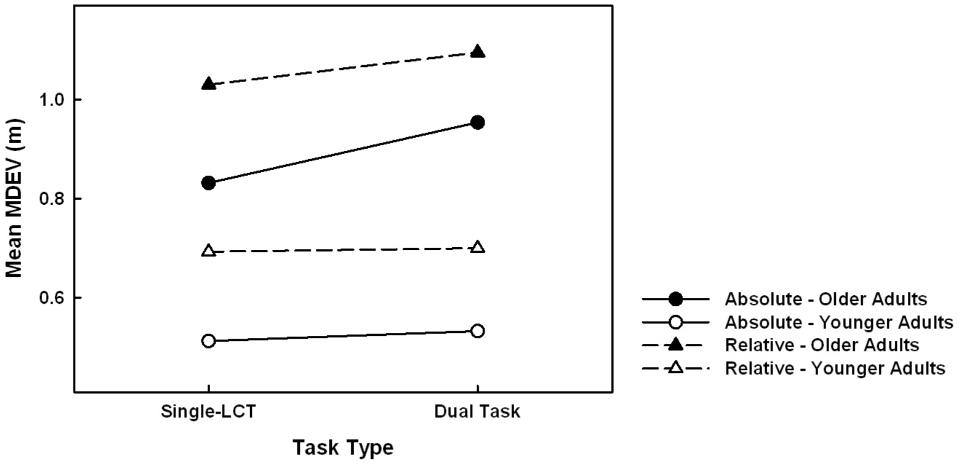


Figure 3.2: Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and age group (younger, older).

for older adults with two separate 1-way (task type) ANOVAs for absolute and relative windows respectively showed that there was a significant simple main effect of task type when using the constant window ($F[1,\ 11] = 13.30$, $\eta_p^2 = .55$, $p < .01$), but, although showing a tendency, not for the relative window ($p = .07$). This means that especially for older adults, the choice of a particular window calculation method, influenced driving performance measures: MDEV values increased significantly in dual-task conditions as compared to single-LCT conditions when using the absolute window.

Figure 3.3 presents mean MDEV as a function of segment and window. The 2-way interaction between those factors was significant ($F[1,\ 21] = 120.64$, $\eta_p^2 = .85$, $p < .001$). A follow-up on this interaction with two separate 1-way (segment) ANOVAs for absolute and relative windows respectively, showed that there was a significant simple main effect of the factor segment on both absolute ($F[1,\ 22] = 102.28$, $\eta_p^2 = .82$, $p < .001$) as well as relative windows ($F[1,\ 22] = 211.98$, $\eta_p^2 = .91$, $p < .001$). The interaction can thus be explained as follows: The calculation method used had no effect on MDEV values in LK segments. In LC segments however, MDEV values are significantly higher with the relative window as compared to the absolute window.

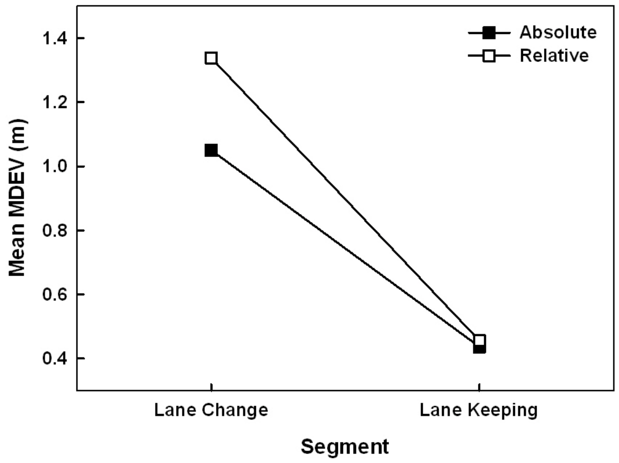


Figure 3.3: Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping) and window (absolute, relative).

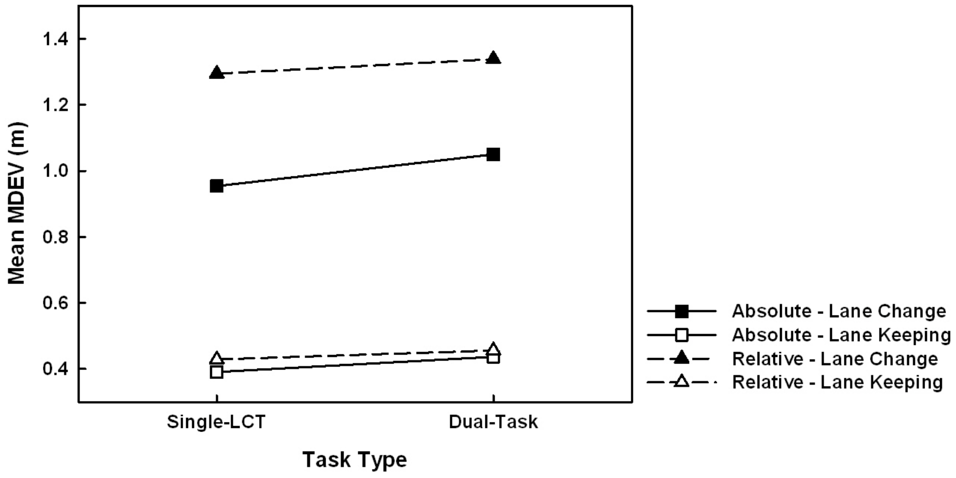


Figure 3.4: Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping).

Figure 3.4 shows mean MDEV as a function of task type, window and segment. The 3-way interation including those factors showed a tendency, but just missed to reach significance ($F[1,\ 21] = 3.54$, $\eta_p^2 = .14$, $p = .07$). It seems to provide however an indication that MDEV values in LK segments hardly differed, independent of task type and window used for calculation. However, for the LC segments, differences were found. First of all, the MDEV values for LC segments were considerably higher when the relative calculation method was used. Second, although small, the increase in MDEV values going from the single-LCT to the dual-task condition seemed to be stronger for absolute windows as compared to relative windows.

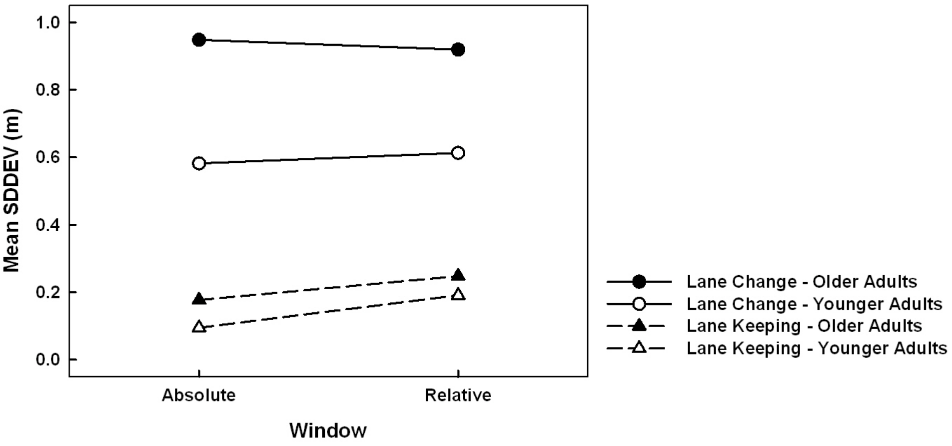


Figure 3.5: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older).

*SDDEV.* Figure 3.5 presents mean SDDEV as a function of window, segment and age group. The ANOVA revealed a main significant effect of window on SDDEV ($F[1,\ 21] = 99.57$, $\eta_p^2 = .83$ $p < .001$): SDDEV in relative windows was higher (M = .49 m) than SDDEV in absolute windows (M = .44 m). There was a significant main effect of age as well ($F[1,\ 21] = 35.07$, $\eta_p^2 = .63$, $p < .001$). Older adults showed more variability (M = .56 m) around the optimal track than younger adults (M = .37 m). Those main effects were however modulated by significant 2-way interactions. The ANOVA revealed a main significant 2-way interaction between the factors window and age on SDDEV ($F[1, 21] = 25.41$, $\eta_p^2 = .55$, $p < .001$) as well as segment and age ($F[1, 21] = 16.51$, $\eta_p^2 = .44$, $p < .01$) and finally window and segment ($F[1, 21] = 89.63$, $\eta_p^2 = .81$, $p < .001$).

All those 2-way interactions were modulated by a significant 3-way (window x segment x age) interaction ($F[1, 21] = 6.46$, $\eta_p^2 = .24$, $p < .05$). To follow up on this interaction, two separate 2-way (window x age) ANOVAs were conducted for LC and LK segments respectively. The interaction turned out to be significant for both LC ($F[1, 21] = 24.96$, $\eta_p^2 = .54$, $p < .001$) as well as LK segments ($F[1, 21] = 4.59$, $\eta_p^2 = .18$, $p < .05$). A further analysis with four separate 1-way (window) ANOVAs for each age group in each driving segment, revealed, for LC segments, a significant simple main effect of window for younger adults ($F[1, 10] = 21.66$, $\eta_p^2 = .68$, $p < .01$) as well as older adults ($F[1, 10] = 5.74$, $\eta_p^2 = .34$, $p < .05$). For LK segments, there was a significant simple main effect of window for both younger ($F[1, 10] = 357.01$, $\eta_p^2 = .97$, $p < .001$) and older adults as well ($F[1, 11] = 100.58$, $\eta_p^2 = .90$, $p < .001$). The 3-way interaction shown in Figure 3.5 can thus be explained as follows: In LC segments, SDDEV values of older adults decrease when relative windows are used, whereas for younger adults, SDDEV values increase with relative windows. In LK segments however, SDDEV values for both age groups increase with the use of relative windows, but this increase is stronger for younger adults than for older adults.

Figure 3.6 presents mean SDDEV as a function of task type and window. Task type interacted significantly with the factor window ($F[1, 21] = 9.11$, $\eta_p^2 = .30$, $p < .01$). To follow up on this interaction two separate 1-way (task type) ANOVAs for each window respectively were conducted. The ANOVA showed that there was no main effect of task type when using the relative window as a basis for calculating SDDEV ($p > .26$), and that this effect only showed a tendency when using the absolute window as a basis for calculation ($F[1, 22] = 3.76$, $\eta_p^2 = .15$, $p = .07$). This might be an indication that the 2-way interaction can be explained by the fact that the increase in SDDEV values going from the single-LCT to the dual-task condition is stronger when absolute windows are used as compared to relative windows.

Figure 3.7 presents mean SDDEV as a function of task type, window and segment.

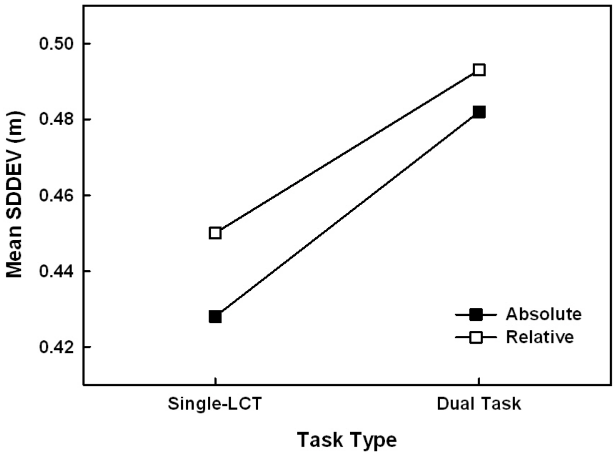


Figure 3.6: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of (a) task type (single-LCT, dual task) and window (absolute, relative).

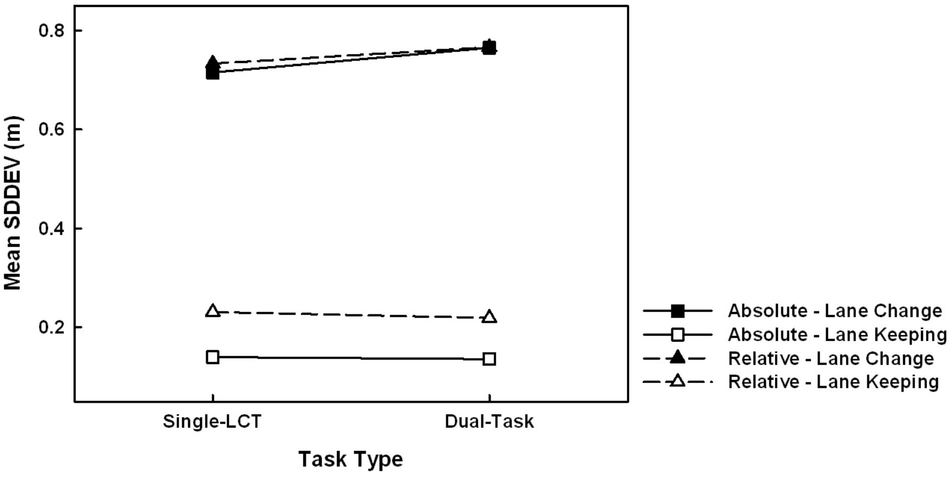


Figure 3.7: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single, dual), window (absolute, relative), and segment (lane change, lane keeping).

The 3-way interaction between those factors just missed to reach significance ($F$[1, 21] = 4.28, $\eta_p^2$ = .17, $p$ = .05) but nevertheless provides an indication that although the use of relative windows yielded slightly higher SDDEV values than the use of absolute windows in LK segments, values hardly differed between single-LCT and dual-task conditions. However, for the LC segments, SDDEV values increased going from the single- to the dual-task driving condition and this effect seemed to be a bit stronger for calculations in which the absolute window was used.

*SDSW.* Figure 3.8 presents mean SDSW as a function of window, segment and age group. The ANOVA revealed a significant main effect of window on SDSW ($F$[1, 21] = 135.50, $\eta_p^2$ = .87, $p$ < .001): SDSW was significantly higher when the relative window calculation method was used (M = 8.56°) as compared to the use of the absolute window calculation (M = 5.99°). There was a significant main effect of age on SDSW as well ($F$[1, 21] = 30.66, $\eta_p^2$ = .59, $p$ < .001): SDSW of younger adults was higher (M = 9.30°) than SDSW of older adults (M = 5.26°).

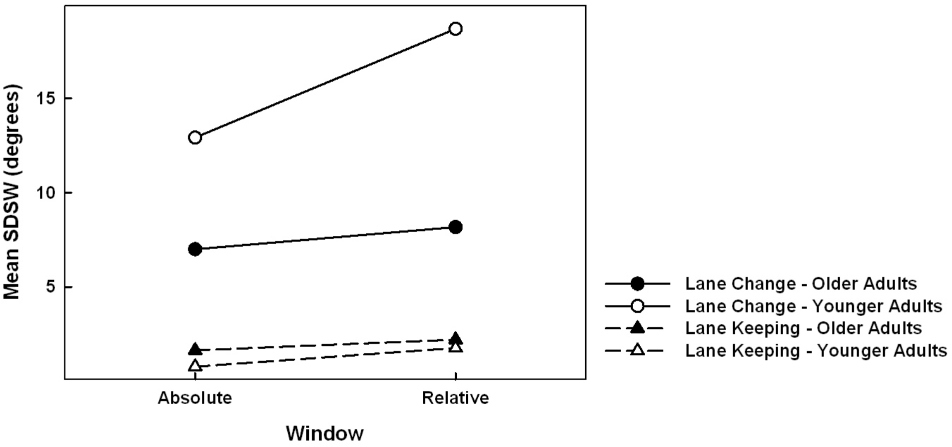


Figure 3.8: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older).

Those main effects were modulated by significant 2-way interactions though. First, the ANOVA revealed a significant interaction between the factors window and age on SDSW ($F[1, 21] = 36.09$, $\eta_p^2 = .69$, $p < .001$). There was furthermore a significant 2-way interaction between segment and age ($F[1, 21] = 49.02$, $\eta_p^2 = .70$, $p < .001$) and finally the 2-way interaction between the factors window and segment turned out to be significant ($F[1, 21] = 78.97$, $\eta_p^2 = .79$, $p < .001$). All these 2-way interactions were modulated by a significant 3-way interaction including all factors (window x segment x age) ($F[1, 21] = 47.24$, $\eta_p^2 = .69$, $p < .001$). To follow-up on this interaction we conducted two separate 2-way (window x age) ANOVAs for LC and LK respectively. The 2-way interaction was significant for both LC ($F[1, 21] = 43.89$, $\eta_p^2 = .68$, $p < .001$) as well as LK segments ($F[1, 21] = 6.76$, $\eta_p^2 = .24$, $p < .05$). To further analyse these interactions, LC as well as LK segments with 4 separate 1-way (window) ANOVAs for each age group and each segment were analyzed more into detail. As to the LC segments, there was a significant simple main effect of window for older adults ($F[1, 11] = 12.55$, $\eta_p^2 = .53$, $p < .01$) as well as younger adults ($F[1, 10] = 137.46$, $\eta_p^2 = .93$, $p < .001$). For the LK segments, a similar pattern was found: A significant simple main effect of window for older ($F[1, 11] = 36.39$, $\eta_p^2 = .77$, $p < .001$) as well as younger adults ($F[1, 10] = 119.39$, $\eta_p^2 = .92$, $p < .001$). The 3-way interaction shown in Figure 3.8 can thus be explained as follows: In LC segments, SDSW values increase when relative windows are used and this increase is strongest for younger adults. For LK segments the same pattern can be found: SDSW values are higher when relative windows are used as compared to the use of absolute windows, and again, this increase is strongest for younger adults.

Figure 3.9 presents mean SDSW as a function of task type, window and segment. The ANOVA yielded a main significant effect of task type on SDSW ($F[1, 21] = 46.32$, $\eta_p^2 = .69$, $p < .001$): Steering-wheel variability was higher in single-LCT conditions (M = 7.92) as compared to dual-task conditions (M = 6.64). The factor task type interacted significantly with the factor window as well ($F[1, 21] = 55.84$, $\eta_p^2 = .73$, $p < .001$) and there was a significant 2-way interaction including the factors task type and segment ($F[1, 21] = 45.57$, $\eta_p^2 = .69$, $p < .001$). These 2-way interactions were however modulated by a significant 3-way (task type x window x segment) interaction ($F[1, 21] = 25.57$, $\eta_p^2 = .55$, $p < .001$). To follow up on this interaction, two separate 2-way (task type x window) interactions for LC and LK segments respectively were conducted. The interaction was significant for LC ($F[1, 22] = 40.15$, $\eta_p^2 = .65$, $p < .001$) as well as LK segments ($F[1, 22]$

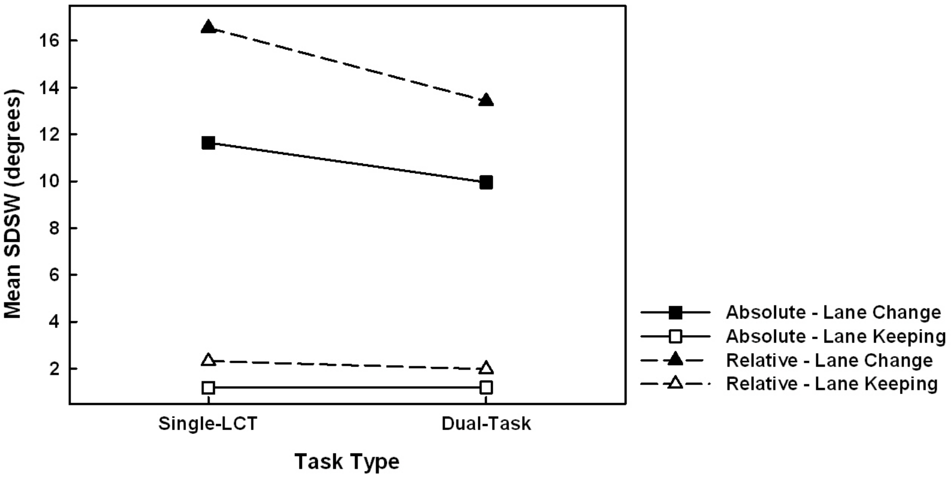


Figure 3.9: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping).

$= 35.08$, $\eta_p^2 = .62$, $p < .001$). A follow-up analysis of the LK segment with two separate 1-way (task type) ANOVAs for each window respectively, revealed that there was no significant simple main effect of task type for absolute windows ($p > .73$), whereas there was a significant simple main effect of task type for relative windows ($F[1, 22] = 8.39$, $\eta_p^2 = .28$, $p < .01$). A follow-up analysis of the LC segment with two separate 1-way (task type) ANOVAs for each window respectively, showed that there was a significant simple main effect of task type for both absolute ($F[1, 22] = 44.77$, $\eta_p^2 = .67$, $p < .001$) as well as relative windows ($F[1, 22] = 48.00$, $\eta_p^2 = .69$, $p < .001$). The observed 3-way interaction in Figure 3.9 can thus be explained as follows: In LK segments, when absolute windows are used, no difference in SDSW values can be observed between the single-LCT and the dual-task condition. However, when relative calculation windows are used for the calculation of SDSW values in LK segments, SDSW values in single-task conditions are higher than those found in dual-task conditions. In LC segments, single-task SDSW values are higher both when using absolute as well as relative windows. The observed difference between tasks however is larger when relative windows are used.

## 3.4 Discussion

The goal of the first research question was to examine whether the use of a relative calculation method can be used to define LC and LK segments. The experiment shows, that the use of numerical derivations to define the duration of each lane change, allow the definition of precise segments (LC and LK) adapted to the actual driving behavior of each participant. It can be assumed that non-confounded (i.e., without including data points from another road segment) driving performance measures can then be extracted from those calculation windows.

The second research question examined the effect of using absolute versus relative LC and LK segments on driving-performance measures in the LCT. The results of this experiment show that the type of calculation window (absolute vs. relative) has a direct effect on the driving performance measures defining lane-change accuracy (i.e., MDEV), lateral-deviation variability while lane-changing (i.e., SDDEV) as well as the amount of steering-wheel movement variability (i.e., SDSW). Values of all those measures were higher using the relative calculation method as compared to the absolute calculation method. This is an indication that the use of relative windows to calculate those values is maybe more closely related to realistic lane-change behavior: As this calculation method allows the calculation of precise windows, adapted to each individual driving style, no confounding data points (e.g., data points from other road segments) are included, yielding performance measures that better reflect individual driving behavior. The use of a relative segment yielded especially a strong effect in LC segments: By a more adequate segment segregation, LK data were excluded from LC data and therewith avoided an artificial underestimation of those driving performance measures.

Finally the third research question examined whether the use of individually adapted segments better reflect age differences in driving behavior. The results in this experiment show that age differences are more pronounced when using relative calculation windows. Lateral-deviation variability (i.e., SDDEV) while lane-changing was higher and steering-wheel movement variability (i.e., SDSW) was lower for younger adults when calculated using the relative windows. An explanation for this finding could be that by using relative windows for the calculation of those measures, the actual LC segment was better adapted to the driving behavior of younger adults: LC segments were shorter and periods in which variability was lower (i.e., LK segments) were not taken into account for the calculation of variability measures. By calculating those values based on actual driving behavior, the effect thus got less attenuated by the inclusion of erroneous data points, causing higher lateral-deviation variability and lower steering-wheel movement variability for younger adults. In other words: By taking absolute windows, values for lateral-deviation variability while lane-changing are potentially underestimated and values for steering-wheel movement variability are potentially overestimated for younger adults.

The data show furthermore that using absolute windows for the calculation of lane-change accuracy values, might actually lead to an overestimation of MDEV values for older adults, especially in dual-task conditions. This finding can again be explained by the rigidity of absolute windows: The critical visual inspection of pilot data from a representative data set including younger and older adults had shown that older adults need in general more time to respond to a traffic sign indicating the lane-change. If on top of that, they are distracted by a secondary task, reactions to traffic signs will potentially be even slower (i.e., lane-change initiation will be slower and lane-change time will increase). By defining the LC segment as an absolute segment, with a fix length and not taking into account the actual lane-change behavior of the participant, lane-change accuracy values will potentially be overestimated.

## 3.5 Conclusion

As a general conclusion, it can be stated that the kinematics of the vehicle's displacement can be used to define the start as well as the length (and therewith the end) of each lane-change maneuver. This method alllows defining relative calculation windows (LC and LK) adapted to each participant. Relative windows, in comparison to absolute windows, seem to be more precise and more representative of actual driving behavior, by taking into account reaction time until the lane change starts, as well as actual lane-change maneuver time. By taking into account individual lane-change behavior, relative windows seem furthermore better adapted to reflect age differences on the LCT. Depending on the factors taken into account, the use of absolute windows can lead to over-or underestimations of certain driving performance measures.

The use of relative windows for the calculation of driving-performance measures has furthermore the advantage of gaining two new variables allowing in-depth insights into lane-change behavior: Reaction-Time until Lane Change (RT-LC), the response time of a participant to start the actual lane change after an LCT traffic sign gets visible, as well as Movement Time (MT-LC), which represents the period of time in which a participant is actually changing of lane (for a more in-depth description of those measures see the Data Analysis section of the General Method: Chapter 2.4).

Based on empirical data in this experiment, relative windows seem to be more reliable for defining LC and LK segments, which is why this method will be used for analysis throughout this thesis.

# 4 Experiment 2: Effects of Feedback on Dual-Task Performance

## 4.1 Introduction

As discussed in previous sections (e.g. see Chapter 3), the Lane-Change Task (LCT; Mattes, 2003) is an easy-to-use tool to study dual-task driving situations. The LCT has also proven to be a valuable tool for the study of learning effects in dual-task driving situations (Chisholm et al., 2008; Petzoldt et al., 2010; Shinar et al., 2005). Pre-tests in the laboratory showed however that despite explicit instructions to consider the LCT as their primary task, shifts in prioritization occurred from the primary LCT to the secondary VST: Participants had a tendency to focus on the secondary task, thereby ignoring the primary driving task, which resulted in poor driving performance measures on the LCT. This observation is consistent with observations by Levy and Pashler (2008), who showed that dual-task interference was caused by participants ignoring explicit experimental instructions to give maximum priority to the driving task, but instead responded to stimuli from the secondary task. If a shift in prioritization between the primary and the secondary task takes place, observed results (including for example learning effects) might not be representative of real-life driving, where it might appear obvious that the driving task receives far more priority than any secondary task, therewith avoiding accidents and collisions (Fuller, 2007). Researchers like Brumby, Salvucci and Howes (2007) and Horrey (2009) successfully showed that drivers were capable of prioritizing either the driving task or a secondary task in a driving-simulator setting, when experimental instructions and feedback were manipulated to prioritize one task or another.

One reason for the shift in priorities in the experimental setup could be the absence of feedback on the LCT. Whereas the VST provided speed and accuracy information at the end of each experimental block, participants did not receive any information about their driving performance. Brumby and colleagues (2007) successfully used feedback as a method to emphasize the focus towards either the primary or the secondary task in a simulated dual-task driving situation, while studying strategic dual-task trade-offs. Their findings showed that feedback could actually be used for changing task priority.

In a recent study (Berthon-Donk et al., 2011), we tested the effect of end-of-block summary feedback (Summary Knowledge of Results = SKR) on performance on the LCT. Fourty young individuals (mean age = 24.0 years, age range = 21-29 years) were randomly assigned to the *feedback* or *no-feedback* condition in a counter-balanced fashion. Each participant had to drive 10 experimental blocks on the LCT without secondary task. Participants in the feedback condition received feedback at the end of each experimental block, whereas the no-feedback group received no feedback. The results showed that SKR feedback had an effect on driving performance in the LCT. When feedback was provided, participants were better at performing lane changes and in LC segments, lateral-deviation variability decreased over blocks, indicating a learning effect. Feedback furthermore had the advantage that participants continued over blocks to perform a certain amount of corrective steering-wheel movements, resulting in better lane-change performance. These findings are in accordance with the literature according to which feedback improves performance due to guidance (Salmoni et al., 1984) and motivation (Vollmeyer & Rheinberg, 2000).

### 4.1.1 Research Questions and Hypotheses

Although the above study provided interesting data as to the usefulness of SKR on driving performance data, as well as the effect of feedback on learning effects in the LCT, some questions still are unanswered: Does driving performance feedback on the LCT in the form of SKR help participants to primarily focus on the driving task? And if yes, do older and younger adults equally benefit from feedback on the LCT? And finally, what is the effect of feedback on learning in dual-task driving conditions? In the experiment that follows, a group of younger and older participants had to perform the LCT with a visual secondary task. Half of the participants received feedback about their LCT performance in the form of SKR and the other half did not receive any feedback about their LCT performance.

The first research question examines whether driving performance feedback on the LCT in the form of SKR actually helps participants to focus primarily on the driving task instead of any secondary tasks. Some evidence in favor of this idea comes from studies by Horrey (2006), who showed that drivers were able to effectively prioritize the appropriate task, with enhanced performance on either the driving or a secondary task when they were explicitly instructed to do so. Brumby and colleagues (2007) were also able to successfully shift the focus of attention to either the driving task or a secondary task by providing feedback. Participants were instructed to focus either on the driving task or the secondary task and were provided with feedback on the focus performance variable (feedback was not provided for the non-focus performance variable). We furthermore know that driving behavior is regulated in accordance with drivers' goals and motives (Summala, 1997, as cited in Dogan, Steg & Delhomme, 2011). Dogan and colleagues (2011) showed with a driving-simulator study, that drivers would always prioritize safety over other goals (e.g. fuel saving or time saving) when interacting with other road users or when confronted with a traffic light. Of course those results cannot be directly transferred to the LCT as there is no interaction with other road users nor any presence of traffic lights, but it provides an indication that priorities may be shifted, depending on accorded importance. As such, we expect drivers who receive feedback on their secondary task as well as their driving performance, will apply prioritization strategies to allow for a better control of attention between both tasks in favor of the driving task, as compared to drivers who only receive feedback regarding their secondary task performance. In terms of performance, we expect the drivers with feedback to show better driving performance, because it is expected they will prioritize the driving task over the secondary task. This prioritization of the driving task might result in more dual-task interference, resulting in a performance loss on the secondary task.

The second research question investigates, whether the effect of feedback on prioritization schemes is age-related. In other words: Do younger and older adults equally benefit from feedback on the LCT? According to the *Attentional Resource Theory* by Kahneman (1973), a deterioration of performance can be observed when resources to one or both tasks are reduced due to an exceeding of the total capacity of the system. Prioritization of one task might limit dual-task interference, by allocating the maximum of resources to that task, but this comes at a cost for the non-prioritised task, which will suffer from more interference, due to a reduced allocation of resources. We know from previous research that aging processes affect attentional resources responsible for the distribution of attention between two or more tasks (Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) as well as their prioritization. For this reason we expect that the combination of explicit instructions (i.e., that the driving task has to be prioritized over the secondary task) and feedback will motivate both younger and older adults to respect prioritization schemes. However, as with age the allocation of attentional resources is reduced, we expect that older adults, under the effect of feedback, allocate less attentional resources to the secondary task and concentrate more on the primary task. For this reason, under the effect of feedback, we expect a greater loss of performance on the secondary task for older adults, as compared to younger adults.

The last research question examines the effect of feedback on learning in dual-task driving conditions. Feedback is known to enhance motivation (Vollmeyer & Rheinberg, 2000) and guidance (Salmoni et al., 1984; Schmidt et al., 1990), which are needed for learning. We furthermore know that feedback in the form of SKR, helps in-depth learning without causing a lot of variability between trials (Berthon-Donk et al., 2011; Schmidt et al., 1990). We therefore expect that the group of drivers who receives feedback about their LCT performance in the form of SKR will show larger improvements over blocks due to enhanced motivation and a more gradual improvement of performance (i.e., variability in performance will be reduced), as compared to the group without feedback. We expect this performance improvement to be observable especially in the driving performance measures, as well as the dual-task measures. However, due to potential conflicts in attention allocation (Kahneman, 1973), performance on the secondary task might remain stable or even decline as compared to single-task performance. This effect is expected to be stronger for older adults (Göthe et al., 2007; Wild-Wall & Falkenstein, 2010).

## 4.2 Method

### 4.2.1 Participants

A total of 52 individuals participated in the experiment. Data of 2 of the participants were discarded due to technical problems. Twenty-four younger participants (mean age = 23.5 years, age range = 20-30 years) and 26 older participants (mean age = 69.5 years, age range = 64-75 years) remained for analysis. Younger drivers reported driving on average 8333 km per year and were in possession of a driver license for a mean period of 5.8 years. Older participants reported driving on average 11923 km per year and their mean period of possession of a driver license was 47.9 years.

### 4.2.2 Design and Procedure

The single-session experimental design was applied as described in the General Method section (Chapter 2.3). However, some particularities for this experiment are described in this section. After filling out some demographic questionnaires as well as undergoing different vision tests (visual acuity test, color blindness test; for an in-depth description of those tests see the section Standardized Tests of the General Method, Chapter 2.2.4), participants were pseudo-randomly (using counterbalancing) assigned to either the *feedback* group or the *no-feedback* group. The feedback group received feedback about their overall performance on the LCT, the no-feedback group did not receive any feedback on their performance on the LCT. Feedback expressed lane-change and lane-keeping performance summarized by the mean deviation (MDEV) from an optimal track, ranging from little deviation (0.3 m) to a lot of deviation (2.5 m; Berthon-Donk et al., 2011). Unlike what was done for the analyses, feedback was computed over the entire track (averaging over LC and LK segments). Participants received furthermore an indication of the number of signs misinterpreted or missed (if any; see the Data Analysis section, Chapter 2.4, for more explanations). Feedback was presented at the end of each block, when the vehicle was standing still, on a separate screen which was positioned on the right of the LCT projection, out of Useful Field Of View (UFOV; Edwards et al., 2006), at a visual position of -10° below the horizon of the driving image and about 40° from the middle of the projection screen on which the LCT was shown.

Including the changes described in the previous paragraph, the experiment consisted of a mixed-factors design including 2 between-participants factors and 3 within-participants factors. To assess the effect of absence or presence of feedback on the LCT, the between-participants factor *feedback condition* was used (feedback, no feedback). To examine a possible influence of age on dual-task priorities, the factor *age* (younger, older) was used as a second between-participants factor. The 3 within-participants factors were *task type* (single, dual), which took into account the difference in single- versus dual-task conditions, *segment* (lane change, lane keeping) which assessed differences in driving task complexity and finally the factor *set size* (9, 16, 25) which reflected differences in secondary task complexity.

### 4.2.3 Data Analysis

The data analysis was divided in two sections: First the effect that a secondary task had on performance measures by comparing single- to dual-task data was analyzed. Then, the dual-task condition was analyzed more in detail. For both sections separately driving, visual-search, dual-task, and subjective data were analyzed separately. The different ANOVAs used for each section are shown in tables below. For a more detailed descriptions of each ANOVA, see Appendix B B.

LCT trials were excluded following the different filter sections described in the Data Analysis section 2.4 of the General Method. For the remaining older adults, the mean percentages of excluded LCT trials were 1.07% (algorithm-error filter), 3.46% (RT-LC filter), 0.98% (MT-LC filter), 0.15% (LK filter), 1.77% (wrong-LC filter) and 0.00% (missed-LC filter), for a total of 7.44% discarded trials in the dual-task condition. The total percentage of discarded trials in the single-LCT condition added up to 1.37% for the older adults. For the remaining younger adults ($N = 24$) 1.39% of all dual-task LCT trials were discarded and 0.60% of the single-LCT trials.

Analysis of VST data was done based on RT-VST, errors and misses. For the analysis of RT-VST, only trials in which a target was present and in which no error occurred were used. Furthermore, all trials in which the participant had answered faster than 200 ms were discarded as they were considered anticipations. As no misses occurred in the VST single-task condition, no ANOVA for the measure misses was performed. Eighty trials were discarded for analysis due to an incomplete data set. For the remaining trials ($N = 8820$), 24.6% was excluded due to errors. The mean percentages of errors in lane-change segments were 7.04% (incorrect responses older adults), 2.64% (incorrect responses younger adults), 4.81% (misses older adults) and 0.23% (misses younger adults), adding up to a total of 14.72%. In lane-keeping segments, the mean percentages of errors added up to a total of 9.87%, including incorrect responses older adults (6.93%), incorrect responses younger adults (1.73%), misses older adults (1.15%) and misses younger adults (0.06%) respectively.

For the VST single-task condition, from a total of 9000 trials, 1468 (16.30%) were discarded due to errors. All errors concerned incorrect responses. Older adults produced most incorrect responses (11.94%).

In the section that follows, tables with different ANOVAs for each performance measure type (driving, visual search task, subjective) will be presented. Tables are divided in ANOVAs taking into account single- and dual-task conditions and ANOVAs looking at dual-task data only. For a more detailed description of these ANOVAs see Appendix B.

### ANOVAs Driving Data

Table 4.1: ANOVAs for driving data taking into account single- and dual-task conditions and ANOVAs looking at dual-task data only.

| Factors | Performance Measures | Remarks |
|---|---|---|
| Single versus Dual-Task Data | | |
| feedback x task type x age | RT-LC, MT-LC | LC segments only |
| feedback x task type x segment x age | MDEV, SDDEV, SDSW | LC segments only |
| Dual-Task Data | | |
| feedback x block x age | RT-LC, MT-LC, IRI | |
| feedback x block x segment x age | MDEV, SDDEV, SDSW | |

### ANOVAs Visual Search Data

Table 4.2: ANOVAs for visual search data taking into account single- and dual-task conditions and ANOVAs looking at dual-task data only.

| Factors | Performance Measures | Remarks |
|---|---|---|
| Single versus Dual-Task Data | | |
| feedback x task type x age | RT-VST, Perc. Incorrect Responses | Averaged over set size |
| Dual-Task Data | | |
| feedback x block x segment x age | RT-VST, Perc. Incorrect Responses, Perc. Misses | Averaged over set size |

*Note.* Perc. = Percentage.

### ANOVAs Subjective Data

Table 4.3: ANOVAs for subjective data taking into account single- and dual-task conditions and ANOVAs looking at dual-task data only.

| Factors | Performance Measures |
|---|---|
| Single versus Dual-Task Data | |
| feedback x task type x age | RSME, NASA-TLX |
| Dual-Task Data | |
| feedback x block x age | RSME |

## 4.3 Results

In this results section, due to the large amount of data, results are presented as follows. First some effects of importance for the experiment, but of less relevance for the research questions are reported ("General Effects"). General effects will only be reported, without providing any statistics or figures. Then, only results of relevance for answering the research questions are presented. An overview of the complete data analysis (including statistical analyses and figures for general effects), can be found in Appendix A. Note that due to the amount of data, figures are only included for significant or close to significant ($p = .08$) main effects or interactions. Close to significant interactions will only be described, without providing any exploratory follow-up tests. If an interaction accounts for answering two or more research questions, the statistical analysis of this interaction will only be done once, when the interaction is encountered for the first time. Later references to that same interaction will simply be referred to and main findings will be resumed, before being interpreted for the research question at hand.

### 4.3.1 General Effects

Increasing the overall task-difficulty by adding a visual search task to the driving task had an effect on all driving measures: Lane changes were slower to initiate, movement times were longer and lane changes became less accurate. Steering-wheel variability also decreased in dual-task situations, indicating that less corrective movements were made in the more difficult task condition. Performing the secondary task in addition to the driving task had an effect on visual search performance as well: Both reaction times as well as the proportion of incorrect responses increased in the more difficult dual-task condition. Subjective measures confirm that dual-task conditions were perceived as more mentally demanding than single-task conditions.

It should be noted that especially older adults suffered from the increase of complexity in the dual-task condition: When a secondary task was added, their lane-change duration increased, and their lane-change maneuvers became less accurate and more variable. Evidence for an added difficulty for older adults in the dual-task condition comes from data on the visual search task as well: When the visual search task was performed in the dual-task condition, the proportion of incorrect responses increased significantly. This increase was more important than the observed increase in the proportion of incorrect responses for younger adults.

Older adults performed less well than younger adults in a general manner: They were slower to initiate lane changes, needed more time to change of lanes and their lane changes were less accurate. They furthermore made less corrective steering-wheel movements than younger adults. Their performance on the visual search task was worse as well: In general their reaction times were higher and they produced more incorrect responses than younger adults.

Lane changes were more difficult than lane keeping for all participants: All classical LCT measures (i.e., MDEV, SDDEV, SDSW) improved when lane keeping as compared to lane changing. As expected, lane-change performance was generally worst in the dual-task condition, due to the increase in general task complexity. Here again, older adults suffered more from increased driving task complexity: Their driving performance suffered significantly in the more difficult lane-change road segments as compared to younger adults. Inherent with observed deteriorations in driving performance measures, their corrective steering-wheel movements did not differ between road segments, whereas the corrective steering-wheel movements of younger adults increased in the more difficult lane-change segments. This might indicate that younger adults, despite the increase in driving complexity, remain capable of correcting their trajectory by an increase of steering-wheel movements, resulting in better driving performance than older adults. Visual search data confirmed the difference in driving task complexity between lane-change and lane-keeping segments as well. Most errors (incorrect responses and misses) were produced in the lane-change segments, indicating that this segment was the most demanding (as witnessed by a drop of performance on the driving task) and as a result had the highest impact on performance in the visual-search task. Older adults suffered most from the increase in driving task complexity: While performing lane changes they produced most misses, which might indicate that the demand from the driving task impeded them from responding on time to the secondary task.

Practice had a beneficial effect on driving performance as well as visual search performance measures. First, participants were faster to initiate a lane change. Second, although older adults were always slower than younger adults to respond to a secondary task after initiating a lane-change, their inter-response interval to a secondary task had a tendency to decrease over blocks which is in fact an indication that with practice older adults better deal with the dual-task situation. As to visual search data, reaction time on the visual search task decreased (with a slight tendency for younger adults to benefit from practice a bit more on this aspect) as well as the number of misses, especially for older adults. This indicates that practice was beneficial for both age groups: Younger adults become faster and older adults, not only respond faster to a visual search task, but answer in a more accurate manner as well.

### 4.3.2 First Research Question

Does driving performance feedback on the LCT in the form of SKR help participants focus primarily on the driving task instead of any secondary tasks?

A first indication that feedback affects task priorization in favor of the driving task, comes from observations concerning lane-change movement time (MT-LC). Figure 4.1 presents MT-LC as a function of feedback condition and age group. A significant interaction between the factors age and feedback condition can be observed ($F[1, 46] = 5.39$, $\eta_p^2 = .11$, $p < .05$): An independant-samples $t$-test for the older adults showed that

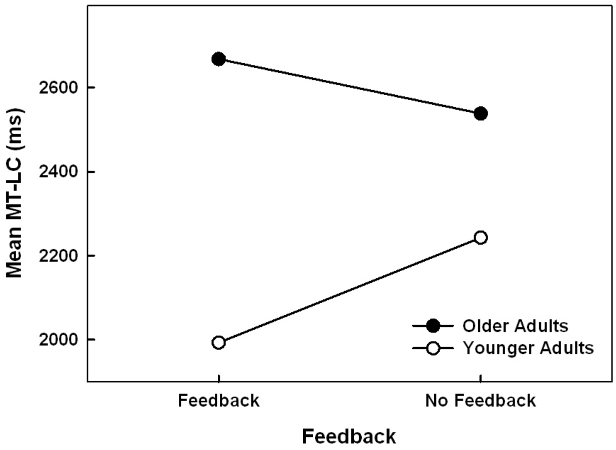


Figure 4.1: Mean Movement Time (MT-LC) as a function of feedback condition (feedback, no feedback) and age group (younger, older).

their movement time did not differ between feedback conditions ($p > .13$), whereas for younger adults, movement times significantly decreased when feedback was provided as compared to the condition without feedback ($t[22] = 2.57$, $p < .01$, one-tailed). This might be an indication that feedback helps especially younger adults to focus primarily on the driving task, therewith enhancing their driving performance measures.

Another line of evidence that feedback has a positive effect on prioritizing the driving task comes from MDEV and (inherent) SDDEV data. Figure 4.2 presents MDEV as a function of segment, age group and feedback condition. The 3-way interaction between those factors was significant ($F[1, 46] = 9.56$, $\eta_p^2 = .17$, $p < .01$). A post-hoc analysis with two separate 2-way (segment x feedback) ANOVAs for younger and older adults respectively, showed a significant 2-way interaction between segment and feedback condition ($F[1, 22] = 5.10$, $\eta_p^2 = .19$, $p < .05$) for younger adults: Independent samples $t$-tests showed that values in the LC segments differed significantly ($t[22] = 1.81$, $p < .05$, one-tailed), whereas values values for the LK segments did not differ statistically between the feedback and the no-feedback condition ($p > .26$, one-tailed). For older adults, the 2-way interaction between segment and feedback was significant as well ($F[1, 22] = 4.36$, $\eta_p^2 = .15$, $p < .05$): Independent samples $t$-tests however showed that there was

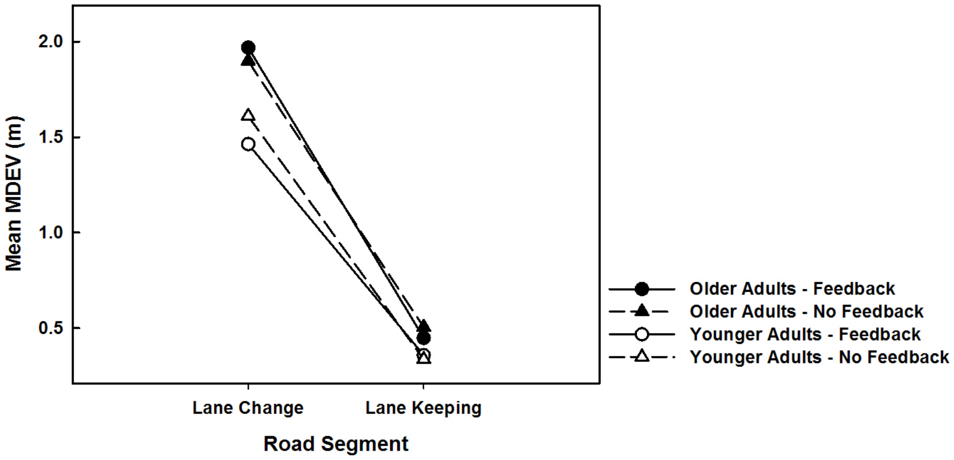


Figure 4.2: Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback).

no difference between LC values ($p > .15$, one-tailed), with and without feedback, and LK values ($p > .06$, one-tailed), with and without feedback. The interaction can thus be explained by the fact that feedback had a positive effect on lane-change accuracy for younger adults in the more difficult LC segments. Feedback had no effect on lane-change accuracy for the older drivers, both in the more difficult LC segments as well as the easier LK segments. This might be an indication that feedback helps younger adults, especially in the more difficult driving segments, to better focus on the driving task and improve their performance on that task accordingly.

Figure 4.3 presents mean MDEV as a function of task type, feedback condition and segment. The 3-way interaction including those three factors showed a tendency but just missed to reach significance ($F[1, 46] = 3.33$, $\eta_p^2 = .07$, $p = .07$). It seems to indicate however that although feedback had hardly any effect in LK segments, it seems to play a role in LC segments. In those more demanding (and more salient!) driving segments, MDEV values increase less going from the single- to the dual-task condition when feedback is provided, than when no feedback is provided. This trend might indicate as well the possible positive effect feedback has on priority management. When the

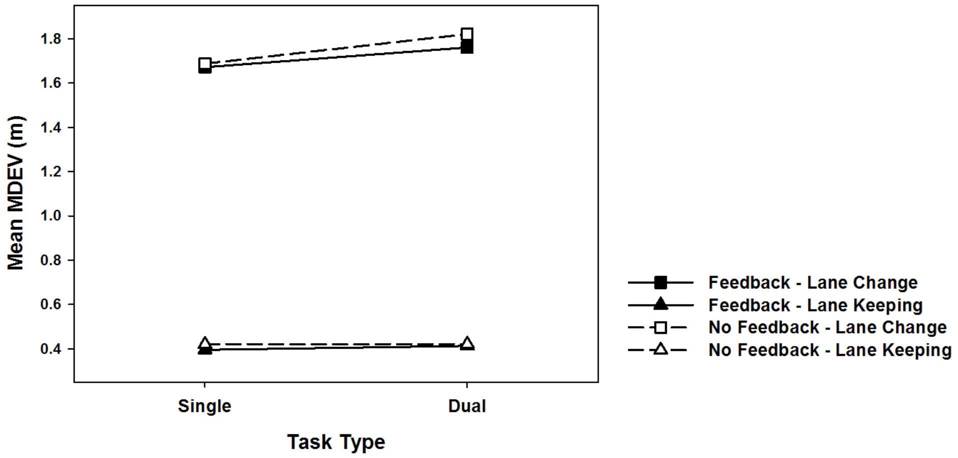


Figure 4.3: Mean Lateral Deviation (MDEV) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping).

driving task is rendered more difficult by a distracting secondary task, priority on the driving task remains, yielding better driving performance measures as compared to the condition without feedback.

Figure 4.4 presents SDDEV as a function of segment, age group and feedback condition. The 3-way interaction between those factors was significant ($F[1, 46] = 8.19$, $\eta_p^2 = .15$, $p < .01$). To further analyze this interaction, two separate 2-way (segment x feedback) ANOVAs for the older and the younger adults respectively were conducted. The 2-way interaction between segment and feedback condition turned out to be significant for younger adults ($F[1, 22] = 4.61$, $\eta_p^2 = .17$, $p < .05$), but, despite a tendency, missed to reach significance for older adults ($p = .08$). Independent-samples $t$-tests on the data of younger adults showed however that SDDEV in the feedback condition just failed to be significant in LC segments ($p = .05$) and was not significant in LK segments ($p = .10$). The interaction shown in Figure 4.4 can thus be explained by the fact that feedback had a strong effect on lane-change variability for younger adults in LC segments only. Again, this might be an indication that feedback helped younger adults to remain focused on the driving task in the more difficult LC segments, leading to lower SDDEV values.

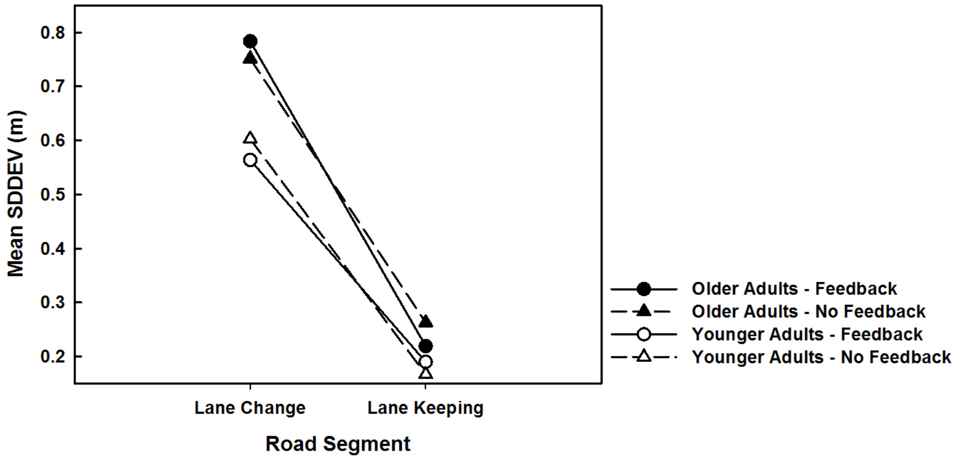


Figure 4.4: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback).

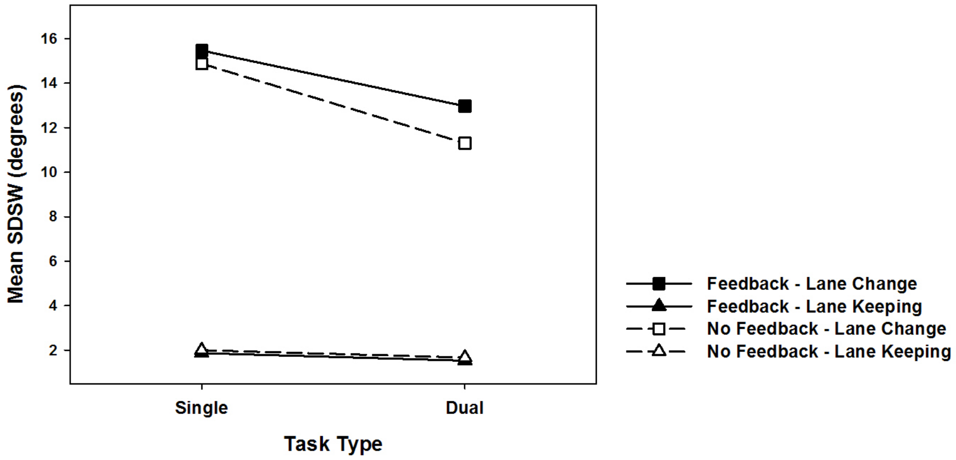
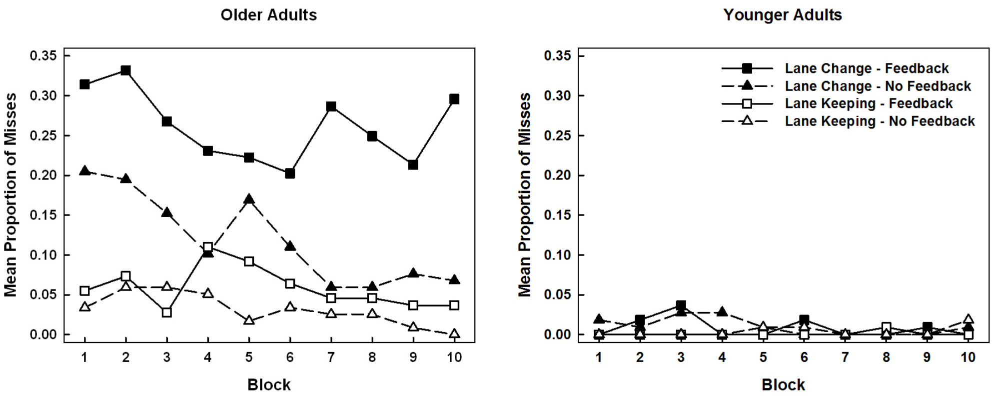


Figure 4.5: Mean Standard Deviation of Steering Wheel angle (SDSW) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping).

Data on the amount of steering-wheel movement variability seem to indicate a positive effect of feedback on task priority as well. Figure 4.5 presents SDSW as a function of task type, feedback condition and segment. The 3-way interaction between those factors was significant ($F[1, 46] = 4.32$, $\eta_p^2 = .09$, $p < .05$). To analyze this 3-way interaction in detail, two separate 2-way (task type x feedback condition) ANOVAs were conducted, for LC and LK segments respectively. The 2-way interaction between task type and feedback did not reach significance for the LK segments ($p > .91$). For the LC segments, this interaction showed a tendency, but just missed to reach significance ($F[1, 48] = 3.49$, $\eta_p^2 = .07$, $p = .07$). This indicates that SDSW has a tendency to decrease in the dual-task condition as compared to the single-task condition, especially when no feedback was provided. For the current research question, this result can be interpreted as follows: In the more difficult driving segments (i.e., LC segments), going from the single-LCT to the dual-task condition resulted in a general decrease in steering-wheel movement variability. However, when feedback was provided, this decrease was less than when no feedback was provided. This might indicate the positive effect of feedback on task priority, helping the drivers remain focused on the driving task (i.e., perform more corrective steering-wheel movements as evidenced by an increased amount of steering-wheel variability; Berthon-Donk et al., 2011).

Looking at visual search data, some indications were found as well, that feedback has a positive effect on priority management in the dual-task condition. Figure 4.6 presents the proportion of misses as a function of block, segment and feedback condition for (a) older adults (b) younger adults. The 3-way interaction including the factors block, segment and feedback condition just missed to reach significance ($F[9, 405] = 2.10$, $\eta_p^2 = .05$, $p = .05$). The 4-way interaction including the factors block, segment, age and feedback condition turned out to be significant however ($F[9, 405] = 2.13$, $\eta_p^2 = .05$, $p < .05$). To analyze this interaction, two separate 3-way ANOVAs including the factors block, segment and feedback condition were conducted for each age group. For the younger adults the 3-way interaction block x segment x feedback condition was not significant ($F[9, 198] = 2.13$, $\eta_p^2 = .05$, $p = .35$). For the older adults, this same interaction was significant ($F[9, 207] = 2.38$, $\eta_p^2 = .09$, $p < .05$). A post-hoc analysis of this interaction with separate 2-way (block x feedback) ANOVAs for each segment, revealed a tendency for both factors to interact in the LC segments ($p = .08$), but no significant interaction



(a) (b)

Figure 4.6: Panel (a): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for older adults. Panel (b): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for younger adults.

could be found in the LK segments ($p > .31$). Those statistics seem to indicate that the interaction shown in Figure 4.6a can be explained as follows: The number of misses in LC segments seems to decrease a bit stronger in the condition without feedback as compared to the condition with feedback for older adults. In the LK segments, feedback is of no influence on learning performance. Feedback thus seems to influence the speed of learning (i.e., the rate at which the number of misses decreases) in the LC segments. When feedback is provided, the number of misses decreases at slower rate than when no feedback is provided. This might indeed indicate that older adults have difficulties focusing on both the LCT and the VST at the same time. When feedback forces them to prioritize the LCT, less attention can be accorded to the VST, resulting in a slower decrease of misses over blocks.

Finally subjective data provide some evidence for the positive effect of feedback on priority management in the dual-task condition. Figure 4.7 presents Mean NASA-TLX ratings for the subscale Temporal Demand as a function of task type and feedback condition. As can be seen, the factor task type showed a tendency to interact with the factor

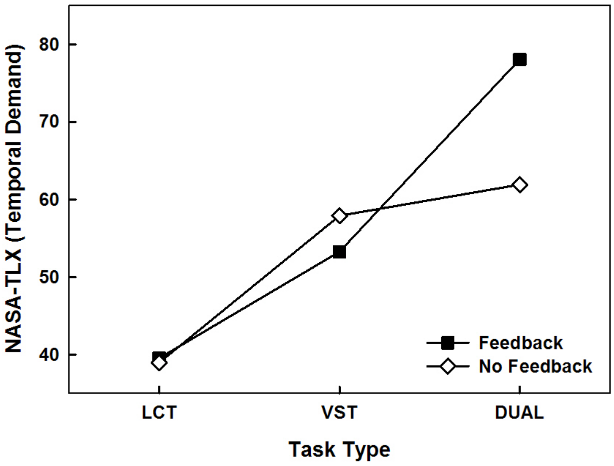


Figure 4.7: Mean NASA-TLX ratings for Temporal Demand as a function of task type (single-LCT, single-VST, dual-task) and feedback condition (feedback, ne feedback).

feedback condition ($F[2, 92] = 3.02$, $\eta_p^2 = .06$, $p = .06$). The ratings of temporal demand seem to increase more or less parallel for the VST in both the FB- as well as the NOFB-condition as compared to the LCT condition. However in the dual-task condition, the subjective rating of temporal demand was much higher for the FB-condition as compared to the NOFB-condition. This indicates not only that participants with feedback felt more stressed by the dual-task condition than participants without feedback, but might furthermore provide an indication that due to "imposed" priority management by feedback, participants had a harder time performing both tasks at the same time (hence rating temporal demand higher).

### 4.3.3 Second Research Question

Is the effect of feedback on prioritization schemes age-related? Or in other words: Do younger and older adults equally benefit from feedback on the LCT?

Looking at driving data, MT-LC data provides an indication that younger and older adults do not equally benefit from feedback on the LCT. Figure 4.1 presents MT-LC as a function of feedback condition and age group. Although the interaction between both factors was significant, a follow-up analysis showed that feedback condition had no effect on movement times for older adults. For younger adults however, movement times significantly decreased when feedback was provided as compared to the condition without feedback. This indicates that feedback has not the same effect on each age group.

Data on lane-change accuracy (i.e., MDEV) as well as the lane-change variability (i.e., SDDEV) provide evidence as well that both age groups don not equally benefit from feedback on the LCT. Figure 4.2 presents MDEV as a function of segment, feedback condition and age group. Although the 3-way interaction between those factors was significant, a post-hoc analysis showed that feedback only had a positive effect on lane-change accuracy for younger adults in the more difficult LC segments. Feedback had no effect on lane-change accuracy for the older drivers, both in the more difficult LC segments as well as the easier LK segments. Figure 4.4 presents SDDEV as a function of segment, feedback condition and age group. And again, although the 3-way interaction between those factors was significant, a follow-up analysis showed that feedback had a positive effect on lane-change variability for younger adults in LC segments only and no effect on lane-change variability of older adults. Again, both results from MDEV and SDDEV data seem to indicate that the effect of feedback on driving performance is age-related.

Other evidence showing that not all age groups equally benefit from the effect of feedback, comes from SDSW data. Figure 4.8 presents SDSW as a function of segment, age group and feedback condition. The observed 2-way interaction between age and feedback condition was significant ($F[1, 46] = 8.65$, $\eta_p^2 = .16$, $p < .01$). Independent-samples $t$-tests showed that SDSW values differed significantly for younger adults between the feedback and the no-feedback condition ($t[22] = 2.51$, $p < .05$, one-tailed). The $t$-test showed however no significant difference between SDSW values in the feedback and the no-feedback condition for older adults ($p = .07$). This indicates again, that the effect of feedback is age-related and that feedback affects SDSW performance for younger adults, but not for older adults. The 3-way interaction between the factors segment, age and feedback shown in Figure 4.8 showed a tendency, but just missed to reach significance

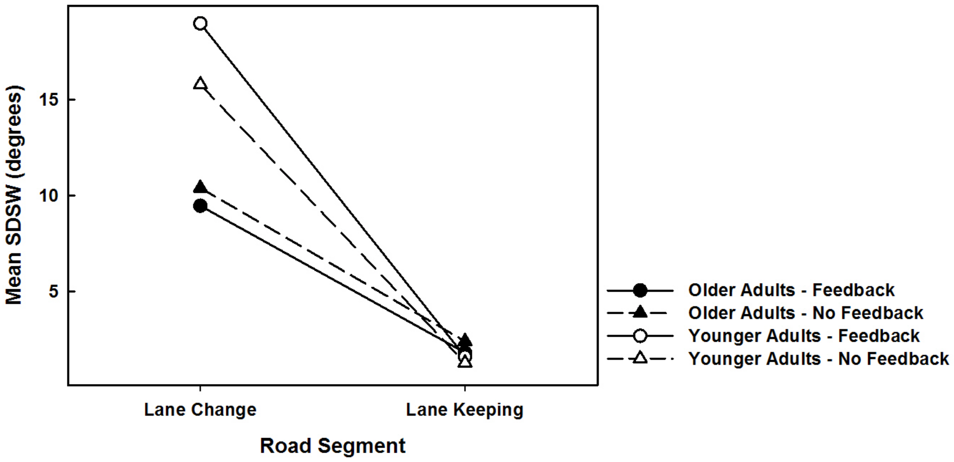


Figure 4.8: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback).

($F[1, 46] = 3.99$, $\eta_p^2 = .08$, $p = .05$). It strongly indicates however that younger adults' SDSW values were higher in LC segments, but especially when feedback was provided. Older adults' SDSW values were higher in LC segments as well, but feedback did not seem to have an effect on those values.

### 4.3.4 Third Research Question

Does feedback affect learning in dual-task driving conditions?

Taking into account the factor block, to assess learning affects on the LCT, some evidence shows that feedback actually has an effect on learning in dual-task driving conditions. Useful data come from the driving measures MDEV and SDSW.

Figure 4.9 presents MDEV as a function of block and feedback condition. There was no significant effect of block ($p = .11$), but there was a tendency for the factors block and feedback condition to interact ($F[9, 414] = 5.95$, $\eta_p^2 = .04$, $p = .08$). Surprisingly,

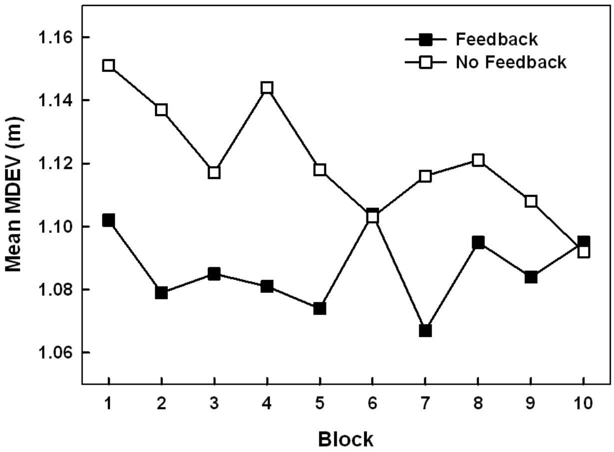


Figure 4.9: Mean Lateral Deviation (MDEV) as a function of block (1-10) and feedback condition (feedback, no feedback).

it seems that participants without feedback improved their MDEV values with practice over 10 blocks, as compared to participants with feedback.

Figure 4.10 presents mean SDSW as a function of block (1-10), segment and feedback condition. The 3-way interaction including all three factors was significant ($F[9,\ 414] = 2.35$, $\eta_p^2 = .05$, $p < .05$). To further analyze this interaction, two 2-way (block x segment) ANOVAs were conducted for the feedback and the no-feedback conditions respectively. The interaction between block and segment was significant for the no-feedback condition ($F[9,\ 216] = 4.72$, $\eta_p^2 = .16$, $p < .01$), but no significant interaction between those factors was found for the feedback condition ($p > .30$). The 3-way interaction can thus be explained by the fact that with practice, feedback had no effect on learning either in the LC and the LK segments. However, when no feedback was presented, the number of corrective steering-wheel movements increased for LC segments, but not for LK segments. Surprisingly again, feedback did not seem to affect learning, and even worse: The group without feedback improved their performance, whereas no effect of practice could be observed for the group with feedback.

Surprisingly, this finding was similar to other interactions including the factors block

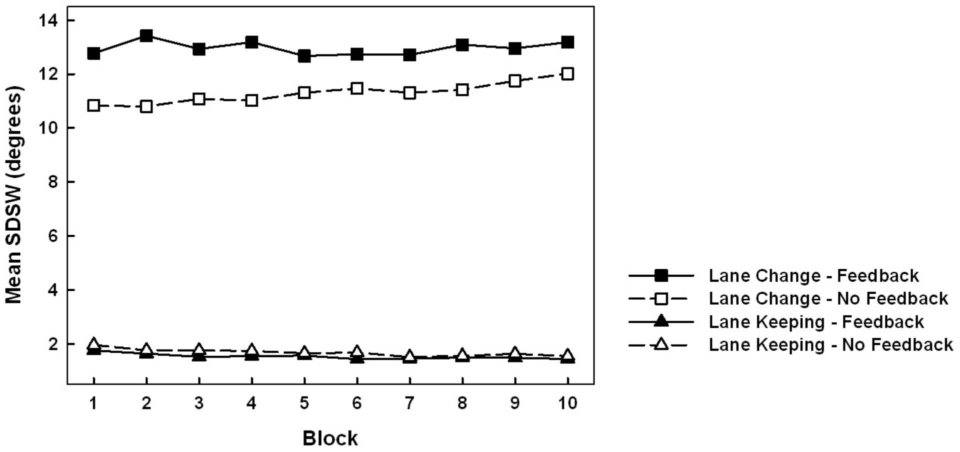


Figure 4.10: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback).

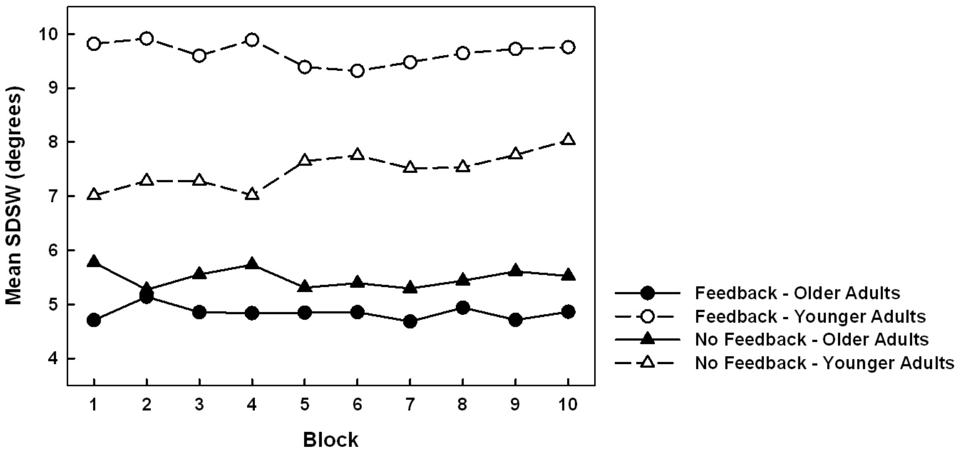


Figure 4.11: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), feedback condition (feedback, no feedback) and age group (young, old).

and feedback condition. Figure 4.11 presents SDSW as a function of block, feedback condition and age group. The 3-way interaction including those factors showed a tendency ($F[9, 414] = 2.09$, $\eta_p^2 = .04$, $p = .08$) but missed to reach significance. It seems to indicate however, that practice had no effect on SDSW performance, except for younger adults receiving, surprisingly, no feedback.

## 4.4 Discussion

As to the first research question, whether feedback helped participants to prioritize the driving task over the secondary task in the LCT, this experiment showed that feedback may have this effect. Several driving performance measures (e.g., lane-change duration, lane-change accuracy and steering-wheel variability) improve under the effect of feedback. This might in fact be an indication that in a dual-task situation, under the effect of feedback, attention was turned in priority to the driving task, hence improving performance on that task. It should be noted however, that feedback was especially beneficial for younger adults and foremost in the more difficult driving segments (we will discuss this more into detail below).

Another indication that feedback has a positive effect on task prioritization comes from data on the visual search task. The number of misses produced by older adults in the more difficult lane-change segments was lower when no feedback was provided, as compared to when feedback was provided. Feedback on the driving task thus worsened performance on the visual search task. Although unexpected, this finding might indicate that feedback helps older participants to focus on the driving task: When interference from a secondary task occurs, to remain focused on the primary task results in performance reductions on the secondary task. These latter findings are in accordance with the attentional resource theory of Kahneman (1973) according to which a drop in performance will be observed when resources for one or both tasks are reduced due to capacity limits of the system. As expected, the effect was stronger for older adults, as due to aging effects, the attentional resources responsible for dividing attention between tasks are less effective (Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010). A final indication for the effect of feedback on priority management in a dual-task condition on the LCT, comes from subjective data: Participants who received feedback on their driving performance reported higher temporal demand (i.e., they felt more time pressure performing their task). This might indicate that the group with feedback, trying to perform equally well on both tasks, felt more pressure than the group without feedback, which allocated more or less attention to one task when demands got high.

These findings are in accordance with findings in previous studies (Brumby et al., 2007; Horrey et al., 2006) in which it was found that explicit oral instructions as well as feedback could change task priorities in a simulated dual-task driving environment. The findings prove furthermore, that even with feedback on both the LCT performance as well as the secondary task performance (i.e., not only on the focus performance variable; Brumby et al., 2007), priorities are respected if participants are explicitly instructed to do so.

The second research question examined whether both older and younger adults benefit equally from feedback on the LCT. The data shows that feedback only affects driving data of younger adults. Feedback does not improve performance for older adults. One explanation for this finding might come from research showing that aging affects sensory-motor functions, resulting in increased reaction times (Alm & Nilsson, 1995; Rubichi et al., 1999; Shanmugaratnam et al., 2010), reductions in flexibility (Davidse, 2004) as well as a reduction in precision of movement (Poston et al., 2009) as well as cognitive functions such as for example executive functions (Adrian et al., 2011). All those age-related changes result in natural performance limitations for older adults as compared to younger adults. The current findings might thus be explained by the fact that older adults are already at their maximum performance, leaving little to no room for improvement, even by adding feedback.

With the third research question it was examined whether feedback had an effect on learning in dual-task driving conditions. Analysis of the current data revealed some unexpected effects concerning feedback and learning. We found that the group of younger adults who did not receive feedback made more accurate lane changes with practice. We found inherently that the group without feedback showed higher steering-wheel variability (which is thought to reflect an increase in corrective steering-wheel movements

Berthon-Donk et al., 2011). In other words, the group of younger adults without feedback performed a larger number of corrective steering-wheel movements, resulting in more accurate lane-change behavior. These findings are not in accordance with previous findings (Berthon-Donk et al., 2011) who found that feedback had a positive effect on learning: In that study, the amount of lane-change variability in lane-change segments decreased with practice. One possible explanation for discrepancies in learning effects under the effect of feedback could be the change in experimental design. In the previous experiment, participants performed 10 experimental LCT blocks only. In the current experiment, not only were the 10 experimental LCT blocks divided in 2 times 5 experimental blocks, but participants furthermore performed a certain amount of experimental VST-blocks as well as experimental blocks of the dual-task condition (for a review of the experimental design, see the section "Design and Procedure" of the General Method, Chapter 2.3). One plausible explanation is that 5 experimental LCT-blocks were too short to replicate any of the effects found in the previous study. Another explanation, especially for the latter 5 experimental LCT-blocks, could be that confounding factors, such as fatigue or boredom, interfered with learning effects.

As to the question why feedback had no effect on learning, whereas learning occurred for the group without feedback, one explanation might be that the group with feedback was so highly motivated, that they very quickly would reach their maximum performance, therewith leaving no room for improvement. Data in favor of this explanation comes from subjective data which showed that younger adults receiving feedback, rated their own performance less severe than those receiving no feedback. This might be due to a motivational effect of feedback (Vollmeyer & Rheinberg, 2000): Participants with feedback could see that their efforts paid off, resulting in potential better performance and providing them an incentive to stay motivated in the next block.

Finally, although the experimental setting of this experiment differed from the previous one (Berthon-Donk et al., 2011) in that it included dual-task data as well as two different age groups, initial findings for the younger age group could partially be replicated. Consistent with the previous study, feedback had a partial positive effect on some of the driving performance measures in the LCT: Lane-change performance variability (as witnessed by SDDEV values) decreased under the effect of feedback. Lane-change accuracy (represented by MDEV values) improved with feedback, especially in LC segments (as compared to higher MDEV values for dual-task conditions without feedback). However, some discrepancies between current findings and findings from the previous study were found as well. In the current study, it was found that feedback had an effect on the amount of steering-wheel movement variability in LC segments, but not in LK segments, whereas in the previous study, feedback was found to have an effect on the steering-wheel variability in LK segments, but not in LC segments (Berthon-Donk et al., 2011). Those discrepancies between both studies might be explained by the difference in data analysis. In the previous study, absolute intervals were used to separate LC from LK segments (Berthon-Donk et al., 2011), whereas in the current study, relative intervals were used to separate LC from LK segments (see Chapter 3 for the rationale behind this choice). For this reason, effects in the previous study might have been over- or underestimated due to the rigidity of each segment, including potential erroneous data, or excluding potential representative data.

## 4.5 Conclusion

All together, it can be said that when the LCT needs to be performed with a secondary task, feedback indeed seems to be a useful tool to focus prioritization strategies on the driving task for both younger and older adults. Feedback has a positive effect on driving performance measures in the LCT as well, although this effect seemed limited to younger adults: Older adults do not seem to benefit from feedback like the younger adults do and their performance does not get better with feedback. Finally, feedback in our experiment did not seem to affect learning in the LCT. On the contrary, although general learning effects were observed with practice, participants without feedback seemed to learn more over blocks than participants with feedback. Further studies will be needed to explain these findings. For the remainder of this thesis however, as this study shows that feedback has a positive effect on priority management in the LCT, SKR feedback will be included for all following studies.

# 5 Experiment 3: Effects of Aging on Learning in Dual-Task Driving Situations

## 5.1 Introduction

The goal of this experiment was to examine the effect of aging on learning in dual-task driving situations. Although some studies in the past tried to focus on parts of this general research question (e.g., Shinar et al., 2005), some particular questions of interest remain to be explored: Does learning have an effect on dual-task driving performance? Does aging affect learning in dual-task driving situations? To answer those questions, an experiment was conducted in which a group of younger and older participants had to perform lane-change maneuvers (Lane Change Test; Mattes, 2003) and a visual search task (Treisman & Gelade, 1980) simultaneously over multiple sessions.

### 5.1.1 Research Questions and Hypotheses

Based on literature reviews, the following research questions were of particular interest: Does learning have an effect on dual-task driving performance and do age-related changes affect learning in dual-task driving situations?

With regard to the first research question, on the basis of Anderson's ACT-R theory (1982), according to which tasks become (partially) automated with practice, thereby freeing up resources, practice was expected to lead to improved performance on both the driving and the secondary task. The Skill-Rule-Knowledge Model (SKR-Model) (Rasmussen, 1983) provides a theoretical framework to explain the potential improvement on both driving as well as secondary task performance measures with practice. This model describes three different levels of cognitive processing that might be used by an individual during task performance. Depending on the nature of the task and the level of experience with the particular situation, a human being operates at one of the three levels. That is, when information is first perceived and interpreted in the processing system, that information is processed cognitively on either the skill-based, knowledge-based or rule-based levels, depending on the individual's degree of experience with the particular situation. The more experience an individual gets, the more information will be processed on a skill-based level, where sensory-motor performances become automated and require very little or no conscious control (Rasmussen, 1983). In the case of driving, it can be expected that with practice, the sensory-motor performance becomes automated, leading not only to improvements on driving performance measures, but also to a freeing of cognitive resources, allowing more attention to be directed to the secondary task, which will lead to improvements on this task as well.

As to improvements in dual-task performance (i.e., performance on both tasks at the same time), Inter-Response Interval (IRI) a variable dedicated to dual-task interference (and described more into detail in the Dual-Task section of the General Method, Chapter 2.4.3) was expected to become shorter with practice. As mentioned above, practice will lead to automation, freeing up resources to better perform both tasks at the same time. These improvements should be reflected in reduced IRI values. Another measure for dual-task performance are eye glance behaviors. Few studies have investigated the effect of learning on glance behavior while driving in a dual-task situation. Popken, Nilsson and Krems (2008) used mean single glance durations to a secondary task display as an indicator for reliance on lane keeping systems. They tested participants in two sessions on the same day and found that mean single glance durations away from the road decreased from the first to the second session. This might reflect a decrease in processing time of complex visual scenes (i.e., to extract the relevant information from the secondary task display) with experience. In line with findings by Popken and colleagues (Popken et al., 2008), drivers are expected to develop a more sophisticated eye scanning behavior over sessions, leading to a decrease in mean glance duration towards the visual search display.

As a complement to the first research question, retention effects will be examined as well. Especially in the context of learning to use IDSS for the driving task, it is of importance to know whether learned skills are readily available even after periods without practice. An extended literature review by Kantak and Winstein (2012) revealed that delayed retention tests (i.e., retention tests that occur after a period of at least 24 hours without practice) more accurately reflect how well the learner had encoded and consolidated a skill for later retrieval. Literature on retention effects on learning in the medical domain (Ackermann, 2009; Gombeski et al., 1982; Kaczorowski et al., 1998) use time-spans of 1 week to several months. They typically find that the largest part of skills is forgotten one week after the last training session. In other words, retention effects can be observed at least 24 hours after the last training session and up till several months after the last training sessions. For this reason, and for logistic reasons, the retention interval took place 3 weeks after the last practice session. And as skill acquisition, as well as retention of this skill, heavily depend on memory processes (Wagner, 2006), which are known to decline at later ages (Borella et al., 2008), performances differences are expected between younger and older adults after this retention period.

The second research question examines whether age-related changes affect learning in dual-task driving situations. As shown by previous studies (for example Caird et al., 2008; Göthe et al., 2007; Verwey et al., 2011) age-related differences are likely to persist, even after extended practice. For this experiment, it is expected that even after practice, older adults' performance will differ from younger adults' performance: Skill acquisition is expected to be slower for older adults than for younger ones, that is, older adults will learn at a slower rate. One reason for this expected difference in learning rate is the age-related decline in cognitive processes (Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) that underlie skill acquisition (e.g., attentional processes, memory functions; Anderson, 1982) and which result in a general slowing of this process. Therefore, age differences in learning are expected to be found on all performance measures: Practice effects will be found, but they will be stronger for younger adults (i.e., younger adults are expected to learn at a faster rate).

## 5.2 Method

### 5.2.1 Participants

A total of 20 individuals (10 younger participants, mean age = 27.6 years, age range = 25-32 years; 10 older participants, mean age = 70.4 years, age range = 67-80 years) participated in the study. Younger drivers reported driving 12000 km per year and were in possession of a driver license for a mean period of 9.3 years. Older participants reported driving 11000 km per year and their mean period of possession of a driver license was 48.6 years.

### 5.2.2 Procedure

To account for effects of practice and learning, the experiment consisted of 4 sessions of approximately 3 hours each. The first three sessions took place within a time frame of 2 weeks with an average inter-session interval of 3 days, the fourth session took place 3 weeks after the third session. Within one session, the single-session experimental design was applied as described in the General Method section (Chapter 2.3). All participants received end-of-block feedback over their driving performance as described in Experiment 2 (Chapter 4: Effect of Feedback on Dual-Task Performance). To allow for the registration of eye movements, electrooculogram (EOG) recording electrodes were attached to the participant's face before starting the first task.

The experiment consisted of a mixed-factors design including 1 between-participants factor and 4 within-participants factors. First the between-participants factor *age* to asses the effect of age on learning effects. Within-participants factors consisted of the number of sessions (*session*, 1-4) examining possible learning effects by practice, the factor *task type* (single, dual) to assess differences in single- versus dual-task conditions, the factor *segment* (lane change, lane keeping) to reflect differences in driving-task complexity and finally the factor *set size* (9, 16, 25) to examine the effect of differences in secondary-task complexity.

### 5.2.3 Data Analysis

For the data analysis, first single- and dual-task data were compared, in order to analyze the effect that adding a secondary task had on performance measures. Second, the dual-task condition was analyzed more into detail. For both sections separately, driving, visual-search, dual-task, and subjective data were analyzed separately. The relevant ANOVAs used for each section are shown in tables below. For a more detailed descriptions of each ANOVA, see Appendix C.

LCT trials were excluded following different exclusion criteria described in the Data Analysis section (Chapter 2.4) of the General Method. Table 5.1 presents the percentage of trials discarded in the single-LCT and the dual-task condition for older and younger adults.

Table 5.1: Percentage of trials discarded in the single-LCT and the dual-task condition for older and younger adults.

| | Alg. Error | RT-LC | MT-LC | LK | Wrong-LC | Missed-LC |
|---|---|---|---|---|---|---|
| Single-LCT | | | | | | |
| Older Adults | 0.20 | 5.39 | 0.84 | 0.54 | 0.29 | 0.00 |
| Younger Adults | 0.00 | 0.45 | 0.63 | 0.33 | 0.10 | 0.00 |
| Dual-Task | | | | | | |
| Older Adults | 1.07 | 4.82 | 2.57 | 0.43 | 1.78 | 0.00 |
| Younger Adults | 0.23 | 0.93 | 0.78 | 0.78 | 0.31 | 0.00 |

*Note.* Alg.Error = Algorithm Error.

For the VST data, only trials in which a correct response was given and where a target had been present were considered for analysis (i.e., half of the trials). From the remaining trials, all trials in which the participant had answered faster than 200 ms were discarded as they were considered anticipations. In the single-VST condition, 0.20% of the remaining trials were discarded for younger adults and 2.52% for older adults. In the dual-task condition, 0.18% of the trials were discarded due to anticipation for younger adults and 5.15% of the trials for older adults.

In the section that follows, tables with relevant ANOVAs for each performance measure type (driving, visual search task, subjective) will be presented. Tables are divided in ANOVAs taking into account single- and dual-task conditions and ANOVAs looking at dual-task data only. For a more detailed description of those measures, see the Data Analysis section in the General Method (Chapter 2.4). For a more detailed description of these ANOVAs see Appendix C.

### ANOVAs Driving Data

Table 5.2: ANOVAs for driving data taking into account single- and dual-task conditions and ANOVAs looking at dual-task data only.

| Factors | Performance Measures | Remarks |
|---|---|---|
| Single versus Dual-Task Data | | |
| session x task type x age | RT-LC, MT-LC | LC segments only |
| session x task type x segment x age | MDEV, SDDEV, SDSW | Averaged over set size |
| Dual-Task Data | | |
| session x set size x age | RT-LC, MT-LC | LC segments only |
| session x set size x segment x age | MDEV, SDDEV, SDSW | |

### ANOVAs Visual Search Data

Table 5.3: ANOVAs for visual search data taking into account single- and dual-task conditions and ANOVAs looking at dual-task data only.

| Factors | Performance Measures | Remarks |
|---|---|---|
| Single versus Dual-Task Data | | |
| session x task type x set size x age | RT-VST, Perc. Incorrect Responses | LC segments only |
| Dual-Task Data | | |
| session x segment x set size x age | RT-VST, Perc. Incorrect Responses, Perc. Incorrect Misses | |

*Note.* Perc. = Percentage.

### ANOVAs Subjective Measures

Table 5.4: ANOVAs for subjective data taking into account single- and dual-task conditions.

| Factors | Performance Measures |
|---|---|
| session x task type x age | RSME, NASA-TLX |

### ANOVAs Dual-Task Measures

Table 5.5: ANOVAs for dual-task measures.

| Factors | Performance Measures |
|---|---|
| session x set size x age | IRI |
| session x segment x set size x age | Vertical Saccade Frequency (VSF),<br>Mean Glance Duration (MGD),<br>Total Glance Duration (TGD) |

## 5.3 Results

In this results section, due to the considerable amount of data and to keep the results section concise, results are presented as follows: First general effects are reported, which are expected based on previous research, but which show that this study is consistent ("General Effects"). General effects will only be reported, without providing any statistics or figures (the according inferential statistics are provided in Appendix C). Then, only results of relevance for answering the research questions are presented. An overview of the complete data analysis (including statistical analyses and figures for general effects), can be found in Appendix C. Note that due to the amount of data, figures are only included for significant or close to significant ($p = .08$) main effects or interactions. Close to significant interactions will only be described, without providing any exploratory follow-up tests. If an interaction accounts for answering two or more research questions, the statistical analysis of this interaction will only be done once, when the interaction is encountered for the first time. Later references to that same interaction will simply be referred to and main findings will be resumed, before being interpreted for the research question at hand. If a significant main effect or a significant simple main effect with the factor *session* occurred, $t$-tests were performed comparing the first and the third session to account for learning effects, and the third and the fourth session to account for retention effects. All $t$-tests were one-tailed.

### 5.3.1 General Effects

Adding a secondary visual search task to the driving task had an effect on almost all performance measures. These findings are as expected and in line with Wickens' four-dimensional multiple-resource model (Wickens, 1984) according to which there will be some interference on the perceptual modality dimension, when two visual tasks are performed at the same time (intra-modal time-sharing). Indeed, creating a dual-task condition by adding a visual search task to the driving task (which hihgly depends on the visual system as well; Owsley & McGwin Jr, 1999), led to performance decreases on both the driving as well as the secondary task as compared to the single-task condition.

Looking into driving performance measures into detail, it was found that lane-change initiations became slower and lane-change movements became longer. Corrective steering-wheel movements decreased when a secondary task was added. As expected, the dual-task situation had an effect on secondary task performance as well: Reaction times and the number of incorrect responses were higher in the dual-task condition as compared to the single-task condition. Subjective estimations of mental effort support the idea that the dual-task condition was more difficult than the single-task condition. All participants rated especially mental, physical and temporal demand higher in the dual-task condition as compared to any of the single-task conditions (single-LCT, single-VST). It should be noted as well, that in the dual-task condition, participants rated the task as demanding more effort and more frustrating, while they rated their own performance as worse. Especially older adults rated the increase in task complexity (by adding a secondary task) as mentally more demanding.

It should be noted however, that depending on the driving segment, task type had potentially opposite effects on lateral deviation variability: In LC segments, lateral deviation variability increased in the dual-task condition as compared to the single-LCT condition, whereas in LK segments, the opposite effect was observed: With increasing task difficulty, variability around the reference track decreased. This is in accordance with findings observed for steering-wheel variability: In LC segments, steering-wheel variability decreased in the dual-task condition as compared to the single-LCT condition. As found by Berthon-Donk and colleagues (Berthon-Donk et al., 2011) a decrease in steering-wheel variability was often paired with a decrease in driving performance.

As expected, this experiment confirmed once more that driving task complexity (i.e., the more difficult lane-change segments versus the easier lane-keeping segments) had an effect on driving performance measures. In the single-task condition, those effects were reflected especially in the ”classical” LCT measures: Lane-change accuracy and variability were higher in LC segments and steering-wheel variability decreased. An in-depth analysis of the dual-task condition showed that an increase in driving task difficulty had an effect on the secondary task performance as well. The number of incorrect responses was significantly higher in LC as compared to LK segments. Especially older adults seem to ”suffer” from an increase in driving task difficulty: The proportion of errors on the visual search task increased stronger in LC as compared to LK segments for older adults.

Driving task complexity was reflected in eye-movement data as well: As expected, glance durations towards the secondary task display were shorter in LC as compared to LK segments. These findings are in line with previous research on glance durations (Chapman, Underwood, & Roberts, 2002; Konstantopoulos, Chapman, & Crundall, 2010; Victor, Harbluk, & Engström, 2005) according to which increasing driving task difficulty (i.e., because of road type, light and weather conditions as well as surrounding traffic complexity) decreased the mean glance duration for glances off the road. At the same time, the number of vertical saccades was higher in the easier LK segments as compared to the more difficult LC segments. These findings are in accordance with findings by Popken and colleagues (2008), who found that when the driving situation became more critical, glance frequencies towards a secondary task display decreased significantly. It thus seems that when driving task difficulty increases, the number of glances off the road as well as their duration decreases. These findings are reflected in a measure combining both glance frequency and duration: Total glance duration. Total glance duration was lower in LC segments as compared to LK segments. In other words: If the driving task gets though, glances are concentrated on the road (which is a good thing!).

As expected, an increase in secondary task difficulty (i.e., more distractor items), increased the mean response time to that task as well as the number of errors produced. There was an effect of age on this finding: An increase in set size resulted in a stronger increase in the proportion of errors produced for older as compared to younger adults. These findings are in accordance with findings from Wilschut and colleagues (2008).

Analyzing dual-task data more into detail, it turns out that secondary task difficulty hardly affected driving performance measures. Lane-change initiation and movement times did not change under different set sizes. Surprisingly, classical LCT measures such as lane-change accuracy (expressed by MDEV), lateral deviation variability (expressed by SDDEV) as well as steering-wheel variability (expressed by SDSW) did not show any effect despite different set sizes either. This might be an indication that indeed, classical LCT measures are not sensitive enough to differences in secondary task difficulty as found by Young, Lenné, and Williamson (2011). According to their research, the LCT is very sensitive to the type of instructions given. If the participants are instructed to imitate a normal driving behavior, the mean deviation is much higher and therefore often less sensitive to changes in cognitive demand. In this experiment, participants were instructed to drive in a way that optimized driving performance measures (i.e., jerking the steering wheel in order to be closest to the reference track). Despite those explicit instructions, differences between older and younger adults were found: Whereas younger adults actually drove in a way that allowed to optimize the MDEV values (i.e., be closest to the reference track), older adults tended to drive in a "smoother" manner, leading to higher MDEV values.

There was an effect of secondary-task complexity on eye-movement behavior as well: With increasing set size, the number of glances towards the secondary task display decreased. However, the mean glance duration increased. These findings are in accordance with other studies (Victor et al., 2005; Wilschut et al., 2008) reporting similar results. An increase in mean glance duration is an indication that indeed a visual scene is more complex, needing more information processing time to analyze it (Underwood, 2007). The authors fail to explain though, why the frequency of glances off the road decreases with increasing set size. One explanation might be that especially older adults do not even try to look for a target on the more difficult search displays. The experimental instructions (i.e., the driving task is the most important task) and the irrelevance of the secondary task for the driving task might have provoked this effect. Evidence for this explanation comes from the high number of misses on the visual search task that older adults produce in dual-task situations, especially for bigger set sizes.

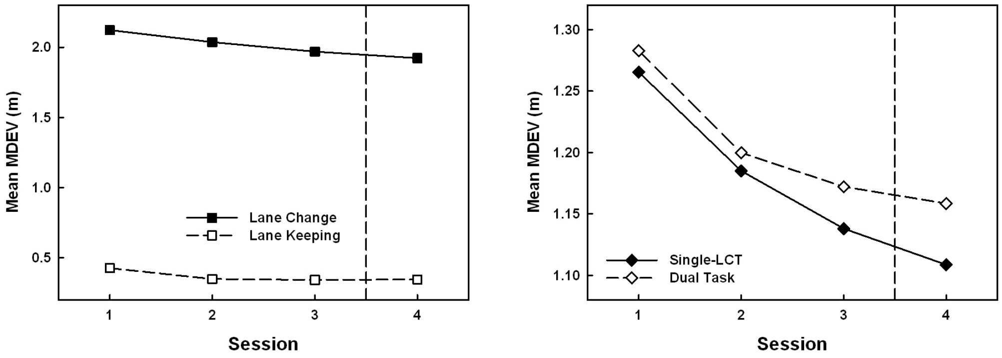
Finally, general age effects could be observed. Younger adults performed more accurate lane changes as witnessed by lower MDEV values. They were also faster to respond to a visual-search task and made less errors. Older adults "suffered" more from an increase in overall task-complexity, by adding a secondary task to the primary driving task, as well: Their IRI values were significantly higher than those of younger adults, indicating that they had more difficulties managing both tasks at the same time. Subjective measures reflected age differences as well: Older adults rated all tasks in general as more demanding and they felt stronger time pressure for all tasks. A look at eye movement behavior revealed that age had an effect on those measures as well: Older adults tended to look down towards the secondary task display (and thus away from the road) less often than younger adults. However, when they looked away from the road, the mean duration of their glances was longer than that of younger adults, indicating that they needed more time to deal with the secondary task. These findings are as expected an in accordance with Hahn and colleagues (2011) according to which older adults need more time to discern relevant from non-relevant information.

### 5.3.2 First Research Question

The first research question examined whether dual-task driving performance changes under the effect of learning. To answer this question, especially main effects of learning and interactions of learning with other factors will be looked at into detail. Retention effects will be looked at as well in this section.

A first indication comes from the onset time of the lane change (RT-LC), which reduces significantly with practice ($F[3,\ 54] = 9.52$, $\eta_p^2 = .35$, $p < .001$). A follow-up analysis of this main effect with paired $t$-tests, showed that learning took place between the first and the third session ($t[19] = 2.56$, $p < .05$), but that the learning effect was not significant between the third and fourth session ($t[19] = 1.91$, $p = .07$). The retention interval thus had statistically no effect on acquired skills.

A similar pattern was found for lane-change duration (MT-LC): The ANOVA revealed a significant main effect of session on MT-LC ($F[3,\ 54] = 9.01$, $\eta_p^2 = .33$, $p < .01$): Practice had a beneficial effect on reducing lane-change duration. Paired $t$-tests showed that there was an effect of learning between the first and the third session on MT-LC ($t[19] = 2.20$, $p < .05$), but, despite a tendency, the difference in performance between the third and fourth session was not significant ($t[19] = 2.06$, $p = .05$). Statistically speaking, the retention period thus did not improve learning, but it did not degrade previously learned skills either.



(a) (b)

Figure 5.1: Mean Lateral Deviation (MDEV) as a function of (a) session (1-4) and segment (lane change, lane keeping) and (b) as a function of session (1-4) and task type (single-LCT, dual-task).

Practice had a positive effect on lane-change accuracy (as witnessed by MDEV) as well. Figure 5.1a presents mean MDEV as a function of session and segment. The ANOVA revealed a significant main effect of session ($F[3,\ 54] = 23.82$, $\eta_p^2 = .57$, $p <$

.001) indicating that practice had a beneficial effect on lane-change accuracy as witnessed by a reduction in MDEV values. The 2-way interaction including the factors session and segment was significant ($F[3,\ 54] = 3.72$, $\eta_p^2 = .17$, $p < .05$) as well. Follow-up tests with separate 1-way (session) ANOVAs for LC and LK segments respectively, revealed a simple main effect of session in LC segments ($F[3,\ 57] = 12.16$, $\eta_p^2 = .39$, $p < .001$) as well as LK segments ($F[3,\ 57] = 9.94$, $\eta_p^2 = .34$, $p < .01$). Although marginal, the interaction seems to come from the fact that practice had a stronger effect on MDEV in LC as compared to LK segments. Those findings are as expected: As LC segments are more difficult than LK segments, participants have more difficulties at first and then improve with practice. To analyze learning effects more into detail, follow-ups on the LC segments with paired-samples $t$-tests were performed: Learning took place between the first and the third session ($t[19] = 3.54$, $p < .01$), but no difference in performance could be observed between the third and fourth session ($p > .20$). The performance in the LK segments showed the same pattern: Learning took place between the first and the third session ($t[19] = 3.45$, $p < .01$), but no difference in performance could be observed between the third and fourth session ($p > .76$). This means that for both segments, the retention interval had no effect on learning. However, performance did not deteriorate either indicating that acquired skills were not forgotten.

Figure 5.1b presents MDEV as a function of session and task type. The 2-way interaction between those two factors turned out to be significant ($F[3,\ 54] = 3.27$, $\eta_p^2 = .15$, $p < .05$). One-way (session) ANOVAs for single-LCT and dual-task respectively revealed a significant simple main effect of session in both the single-LCT condition ($F[3,\ 57] = 22.48$, $\eta_p^2 = .54$, $p < .001$) as well as the dual-task condition ($F[3,\ 57] = 19.59$, $\eta_p^2 = .51$, $p < .001$). The interaction comes from the fact that with practice MDEV decreased at a faster rate in single-LCT conditions as compared to dual-task conditions. This finding is as expected, as the single-LCT condition is known to be easier than the dual-task condition. An in-depth look at learning effects by analyzing the factor session with paired-samples $t$-tests, revealed that in the single-LCT condition, learning took place between the first and the third session ($t[19] = 5.44$, $p < .001$), but that the retention interval had no effect on performance measures ($p > .17$). The same pattern was observed for the dual-task condition: Learning took place between the first and third session ($t[19] = 5.38$, $p < .001$), but not between the third and fourth session ($p > .41$).

This indicates acquired skills remain more or less at the same level, even after a longer period without practice.

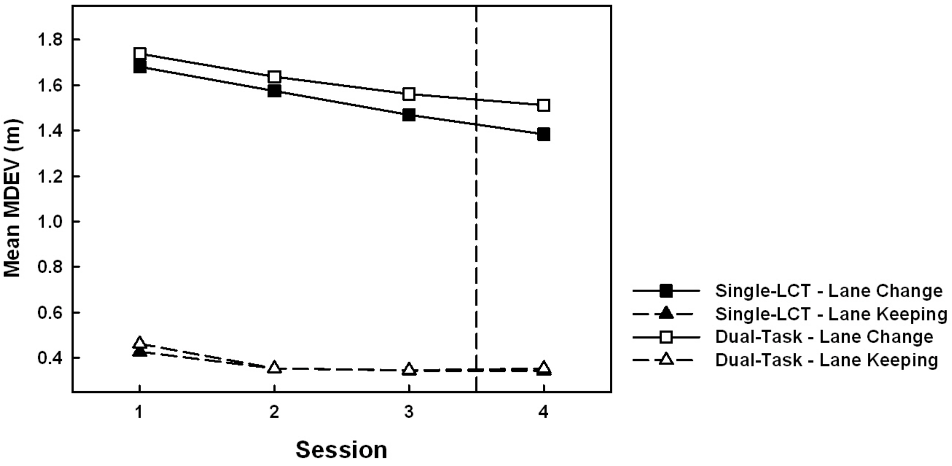


Figure 5.2: Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping).

The 3-way interaction including session, task type and segment, graphically represented in Figure 5.2, showed a tendency but missed to reach significance ($F[3, 54] = 3.08$ $\eta_p^2 = .15$, $p = .06$). It might confirm the findings above though in that learning took foremost place in the more difficult LC segments and the easier single-LCT condition.

Further indications that practice had a beneficial effect on performance measures comes from SDDEV values. Figure 5.3 presents mean SDDEV as a function of session and task type. The 2-way interaction between those factors was significant ($F[3, 54] = 5.82$, $\eta_p^2 = .24$, $p < .01$). A follow-up of this interaction with two separate 1-way (session) ANOVAs for each task type respectively, revealed a significant simple main effect of session in single-LCT conditions ($F[3, 57] = 8.67$, $\eta_p^2 = .31$, $p < .001$), as well as dual-task conditions ($F[3, 57] = 3.46$, $\eta_p^2 = .15$, $p < .05$). However, as can be seen in Figure 5.3, SDDEV values decreased at a higher rate in the single-LCT condition as compared to the dual-task condition, which is again an indication that learning is "easier" in the less difficult single-LCT condition. An in-depth analysis of the learning effect for both conditions with paired-samples $t$-tests comparing the first and the third session, revealed that learning took place in the single-LCT condition ($t[19] = 3.12$, $p < .01$, ), but not in the dual-task condition ($t[19] = 1.97$, $p = .06$). The significant main effect of learning for the dual-task condition revealed by the 1-way (session) ANOVA, might be due to the inclusion of the fourth session in the initial ANOVA. The retention interval (analyzed by comparing the third and fourth session with paired-samples $t$-tests) had no effect on either conditions (all $p$s $> .18$). This indicates that the retention effect did not improve performance, but that acquired skills were not forgotten either.

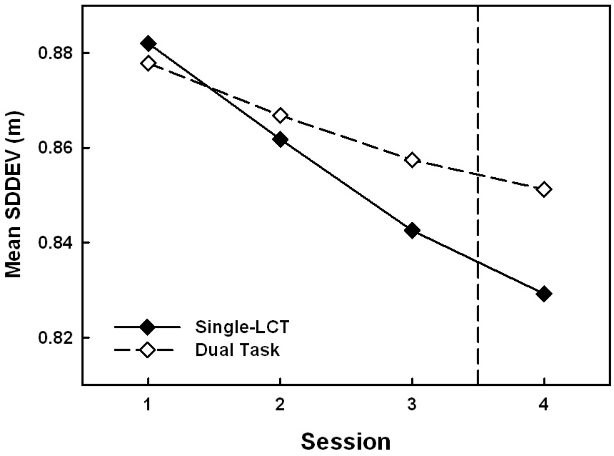


Figure 5.3: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4) and task type (single-LCT, dual).

A final indication for the effect of practice on driving performance comes from steering-wheel variability measures (SDSW). Figure 5.4 presents SDSW as a function of session, task type and segment. The ANOVA revealed a main effect of session ($F[3, 54] = 8.03$, $\eta_p^2 = .31$, $p < .01$): With practice, SDSW values increased. Although not directly visible in figure 5.4, the 2-way interaction between session and segment was significant ($F[3, 54] = 10.48$, $\eta_p^2 = .37$, $p < .001$) just like the 2-way interaction between session and task type ($F[3, 54] = 4.42$, $\eta_p^2 = .20$, $p < .05$) as well as the 2-way interaction between task type and segment ($F[1, 18] = 22.31$, $\eta_p^2 = .55$, $p < .001$). All these 2-way interactions were however modulated by a significant 3-way interaction including the factors

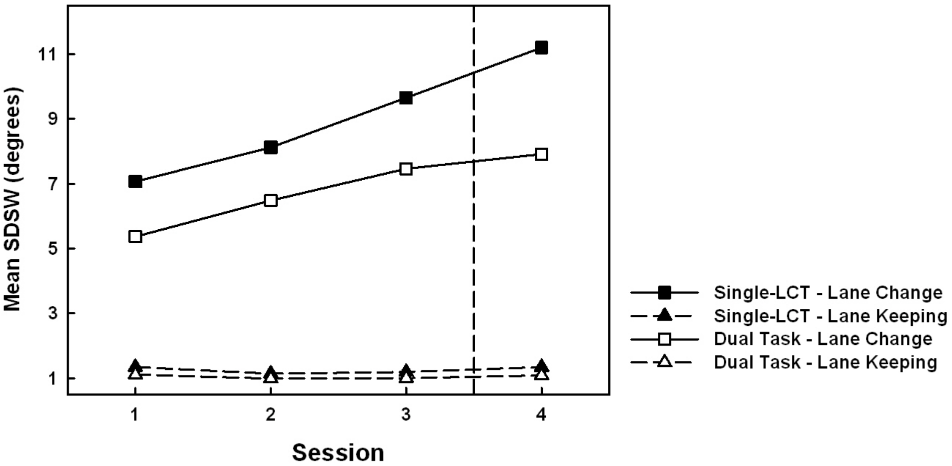


Figure 5.4: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping).

session, task type and segment ($F[3,\ 54] = 4.89$, $\eta_p^2 = .21$, $p < .05$). To follow up on this interaction, we conducted two separate 2-way (session x task type) ANOVAs for LC and LK segments respectively. There was no significant interaction between the factors session and task type in LK segments ($p > .39$), but the interaction reached significance in LC segments ($F[3,\ 57] = 4.79$, $\eta_p^2 = .20$, $p < .05$). Separate 1-way (session) ANOVAs for single-LCT and dual-task conditions respectively showed that there was both a significant simple main effect of session in the single-LCT condition ($F[3,\ 57] = 7.67$, $\eta_p^2 = .29$, $p < .01$) as well as the dual-task condition ($F[3,\ 57] = 6.45$, $\eta_p^2 = .25$, $p < .01$). As can be seen in Figure 5.4, practice had a stronger effect on single-LCT conditions as compared to dual-task conditions. This is in accordance with MDEV and SDDEV results, as generally, when steering-wheel variability increases, driving performance improve as well (Berthon-Donk et al., 2011). This effect was stronger in the easier single-LCT condition, as compared to the more difficult dual-task condition. An in-depth analysis of the learning effects for each task condition with paired-samples $t$-tests, showed that no learning took place between the first and the third session in the dual-task condition ($t[19] = 2.26$, $p < .05$) and only showed a tendency in the single-LCT condition ($t[19] = 2.04$, $p = .06$). This indicates that looking at the first three sessions was insufficient to find any effects. However, if the fourth (retention) session is taken into account, learning takes place. The retention interval only showed a tendency in the single-LCT condition ($p = .09$), but no significant effect in the dual-task condition.

Visual search performance measures provided some indications for the positive effect of practice on performance as well. As to reaction time on the VST (RT-VST), the ANOVA revealed a main effect of session ($F[3,\ 51] = 59.75$, $\eta_p^2 = .78$, $p < .001$): Practice had a beneficial effect on reducing RT-VST values. A follow-up on the learning and retention effects with paired-samples $t$-tests showed that learning took place between the first and the third session ($t[19] = 9.97$, $p < .001$), but that the retention period had no effect on learning (RT-VST values between the third and fourth session did not differ) ($p > .59$).

The proportion of incorrect responses and misses on the VST provided indications for the effect of practice on performance measures as well. The ANOVA revealed a main effect of session on the proportion of incorrect responses ($F[3,\ 51] = 29.55$, $\eta_p^2 = .64$, $p < .001$): The proportion of incorrect responses on the visual search task reduced with practice. A follow-up analysis of this interaction showed that the proportion of incorrect responses significantly decreased between the first and the third session, indicating a learning effect ($t[19] = 7.93$, $p < .001$). The retention period however, had no effect on the proportion of errors produced ($p > .40$).

Figure 5.5 presents the number of misses as a function of session, segment and age group. The ANOVA revealed a significant main effect of session on the proportion of misses ($F[3,\ 51] = 19.78$, $\eta_p^2 = .54$, $p < .001$): Practice had a positive effect on the proportion of misses produced. The ANOVA revealed significant interactions between the factors segment and age ($F[1,\ 17] = 6.42$, $\eta_p^2 = .27$, $p < .05$), the factors session and age ($F[3,\ 51] = 10.02$, $\eta_p^2 = .37$, $p < .001$) as well as the factors session and segment ($F[3,\ 51] = 4.91$, $\eta_p^2 = .22$, $p < .05$). All those 2-way interactions were however modulated by a significant 3-way interaction including those factors (session x segment x age) ($F[3,\ 51] = 3.55$, $\eta_p^2 = .17$, $p < .05$). To further analyze this interaction, separate 2-way (session x segment) ANOVAs were conducted for younger and older adults respectively. There was no significant 2-way interaction between the factors segment and session for

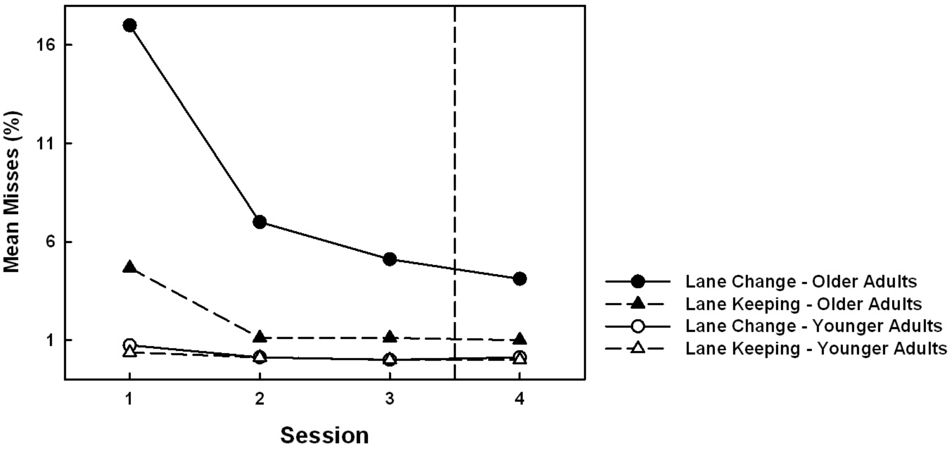


Figure 5.5: Proportion of misses as a function of session (1-4), segment (lane change, lane keeping) and age group (younger, older).

younger adults ($p > .55$), but the interaction reached significance for older adults ($F$[3, 27] = 5.26, $\eta_p^2 = .37$, $p < .05$). A follow-up of this interaction for older adults with two separate one-way (session) ANOVAs for LC and LK segments respectively, showed a significant simple main effect of session for LC segments ($F$[3, 27] = 15.99, $\eta_p^2 = .64$, $p < .001$) as well as LK segments ($F$[3, 27] = 8.31, $\eta_p^2 = .48$, $p < .01$). The interaction shown in Figure 5.5 can thus be explained as follows: The proportion of misses does not reduce for younger adults with practice (in both LC and LK segments). However, older adults benefit from practice for both segments. This learning effect is stronger for LC segments, in which learning takes place at a faster rate. To further analyze these learning effects, paired-samples $t$-tests were performed between the first and third session. They revealed that learning took place in both LC segments ($t$[19] = 4.06, $p < .01$) as well as LK segments ($t$[19] = 3.65, $p < .01$). The retention interval (analyzed with paired-samples $t$-tests between the third and fourth session) showed no significant effect for both LC ($p > .77$) as well as LK segments ($p > .74$). This means that the number of misses remained the same despite a longer period without practice.

Figure 5.6 presents the proportion of misses as a function of session, age group and set size. The factor session interacted significantly with the factor set size ($F$[6, 102] =

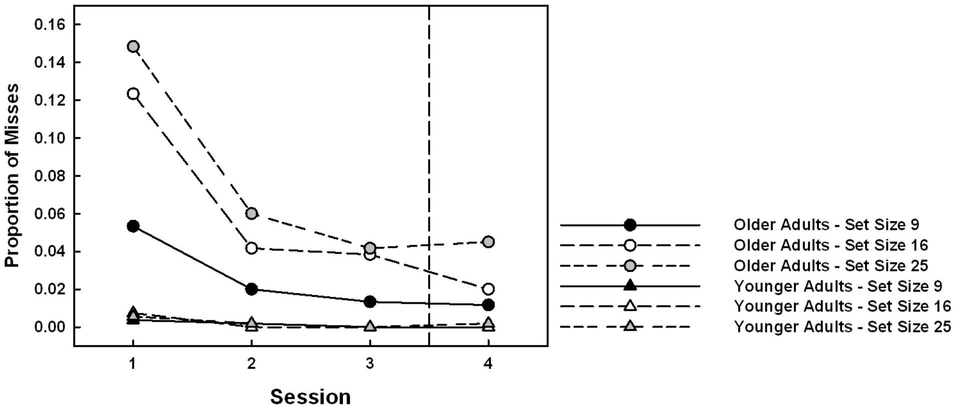


Figure 5.6: Proportion of misses as a function of session (1-4), age group (younger, older) and set size (9, 16, 25).

3.48, $\eta_p^2$ = .17, $p < .01$). Three separate 1-way (session) ANOVAs for the set sizes 9, 16 and 25 respectively showed that there was a significant simple main effect of session for set size 9 ($F$[3, 54] = 5.90, $\eta_p^2$ = .25, $p < .05$), set size 16 ($F$[3, 54] = 10.80, $\eta_p^2$ = .38, $p < .001$) as well as set size 25 ($F$[3, 54] = 13.34, $\eta_p^2$ = .43, $p < .001$). The interaction seems to come from the differences in learning rate with practice: Learning was strongest for the VST-display with 25 items, followed by the display with 16 items and finally the display with 9 items. The 25-item display was maybe the most difficult, therewith leaving most potential for learning.

Subjective measures provided evidence for the effect of practice on performance as well. Figure 5.7 shows mean RSME as a function of session and task type. The ANOVA yielded a significant main effect of session ($F$[3, 54] = 12.53, $\eta_p^2$ = .41, $p < .001$): With practice over sessions the subjective rating of mental demand decreased for all tasks. The 2-way interaction between session and task type was significant ($F$[6, 108] = 4.53, $\eta_p^2$ = .20, p < .01): With practice, all tasks were rated less mentally demanding, but the decrease was strongest for the dual-task condition. This indicates that participants "suffered" most from this condition at first, but that practice was potentially beneficial for reducing task demand, leading to a deterioration in subjective ratings.

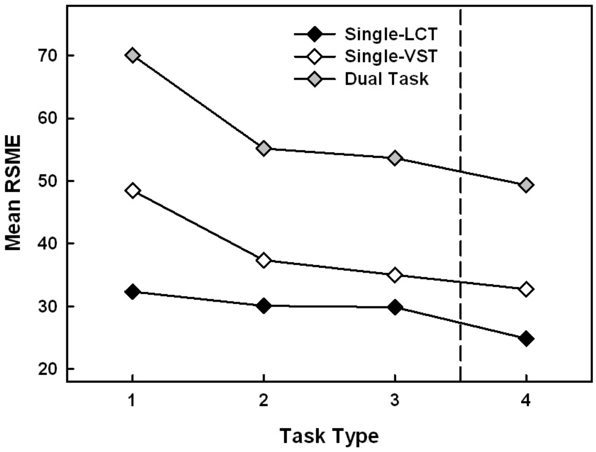


Figure 5.7: Mean Rating Scale of Mental Effort (RSME) as a function of session (1-4) and task type (single-LCT, single-VST, DUAL).

Table 5.6: Results of 3-way (session x task type x age) mixed-factors ANOVAs performed on each of the NASA-TLX subscale ratings. Presented are the NASA-TLX subscales for which a significant main effect of session was found, the factors that reached significance or close-to-significance on each of those scales, the degrees of freedom (df), $F$-values, effect sizes ($\eta_p^2$) and probability levels for statistical significance ($p$).

| NASA-TLX Subscale | Factor | df | $F$ | $\eta_p^2$ | $p$ |
|---|---|---|---|---|---|
| Mental Demand | Session | 3,54 | 12.48 | .41 | $< .001$* |
| Temporal Demand | Session | 3,54 | 4.24 | .19 | $< .05$* |
| Effort | Session | 3,54 | 5.50 | .23 | $< .01$* |
| Frustration | Session | 3,54 | 3.94 | .18 | $< .05$* |

Table 5.6 presents the main effect of session as a result of the 3-way mixed-factors ANOVA on some of the 6 subscales of the NASA-TLX. The factor session had a significant effect on ratings of mental demand, temporal demand, effort and frustration: Practice significantly reduced subjective ratings for those aspects of mental task demand.

Of main interest for this research question are the dual-task measure IRI and eye-movement measures. The ANOVA revealed a significant main effect of session on IRI ($F[3, 51] = 4.56$, $\eta_p^2 = .21$, $p < .05$): With practice IRI decreased. A follow-up analysis with paired-samples $t$-tests for the first and third session showed a learning effect ($t[19] = 2.65$, $p < .05$): IRIs significantly decreased with practice. The analysis between the third and fourth session however showed that the retention interval had no effect on IRI data ($p > .20$).

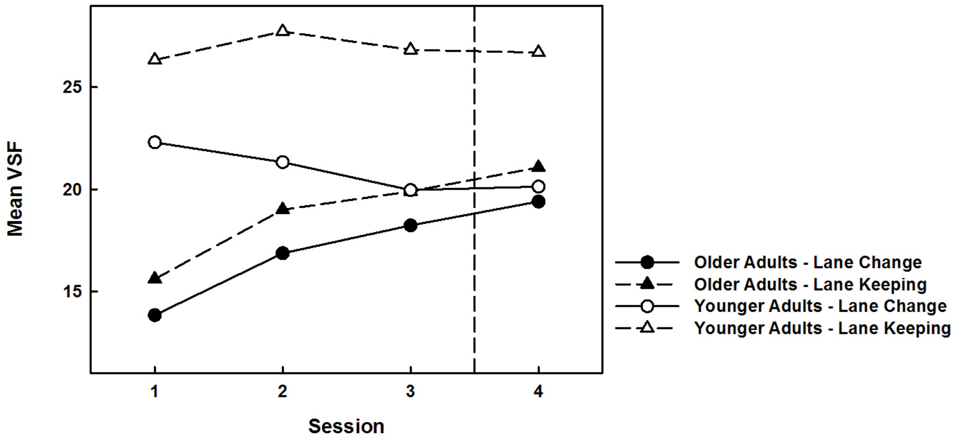


Figure 5.8: Mean vertical saccade frequency (VSF) as a function of session (1-4), age group (younger, older) and segment (lane change, lane keeping).

Eye-movement measures provide some indications for the effect of practice on performance as well. Figure 5.8 presents VSF as a function of session, age and segment. There was no main effect of session, which indicates that the number of display glances did not change with practice. The interaction between session and age was significant though ($F[1, 108] = 5.07$, $\eta_p^2 = .22$, $p < .05$). To follow up on this ANOVA, two separate 1-way (session) ANOVAs were conducted, for younger and older adults respectively. The factor session had no effect on younger adults ($p > .66$) indicating that their vertical saccade behavior did not change with practice. However, there was a significant simple main effect of session for older adults ($F[3,\ 27] = 9.59$, $\eta_p^2 = .52$, $p < .01$), which indicated that with practice, VSF for older adults increased. This might reflect that with practice, older adults feel more confident while performing the task and take their eyes more off the road to look at the display more often. A follow-up analysis of this learning effect for older adults with paired-samples $t$-tests showed that there was a learning effect which took place between the first and third session ($t[9] = 2.93$, $p < .05$), but revealed no effect of the retention interval on VSF ($p > .18$).

The interaction between session and segment showed a tendency, but did not reach significance ($F[1,\ 108] = 3.45$, $\eta_p^2 = .16$, $p = .08$). It seems to indicate though that with practice, VSF in the LK segments increases more than in the LC segments. This finding, although a tendency only, is as expected: As the LK segments are easier, participants feel more confident taking their eyes off the road, which might have led to an increase in VSF.

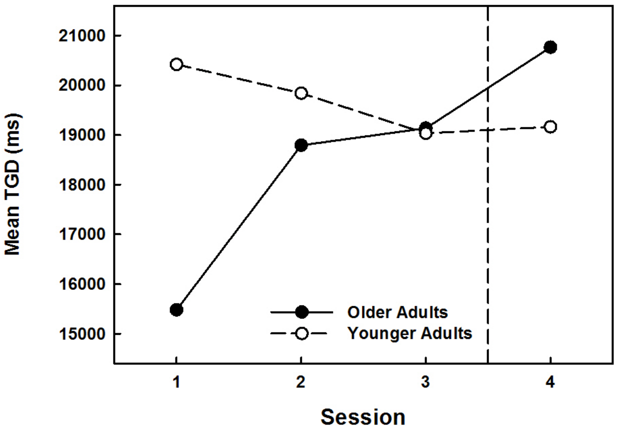


Figure 5.9: Total Glance Duration (TGD) as a function of session (1-4) and age group (young, old).

Figure 5.9 presents mean TGD as a function of session and age group. The interaction between session and age group was significant ($F[1,\ 108] = 4.74$, $\eta_p^2 = .21$, $p < .05$).

To follow up on this interaction separate 1-way (session) ANOVAs were conducted for younger and older adults respectively. There was no significant simple main effect of session on TGD for younger adults ($p > .55$), but there was a significant simple main effect of session for older adults ($F[3, 27] = 6.38$, $\eta_p^2 = .42$, $p < .05$) indicating that the TGD of older adults increased significantly with practice. A follow-up analysis on this effect of session with paired-samples $t$-tests showed that the simple main effect actually came from a significant learning effect between the first and second session only ($t[9] = 2.61$, $p < .05$). The comparison between the first and the third session showed a tendency ($t[9] = 2.13$, $p = .06$), but did not reach significance. It indicates however that adding an extra session has a potential effect on TGD. The comparison between the third an fourth session, the retention period, showed a tendency as well, but missed to reach significance ($t[9] = 2.17$, $p = .06$).

### 5.3.3 Second Research Question

The second research question examines whether age-related changes affect learning in dual-task driving situations.

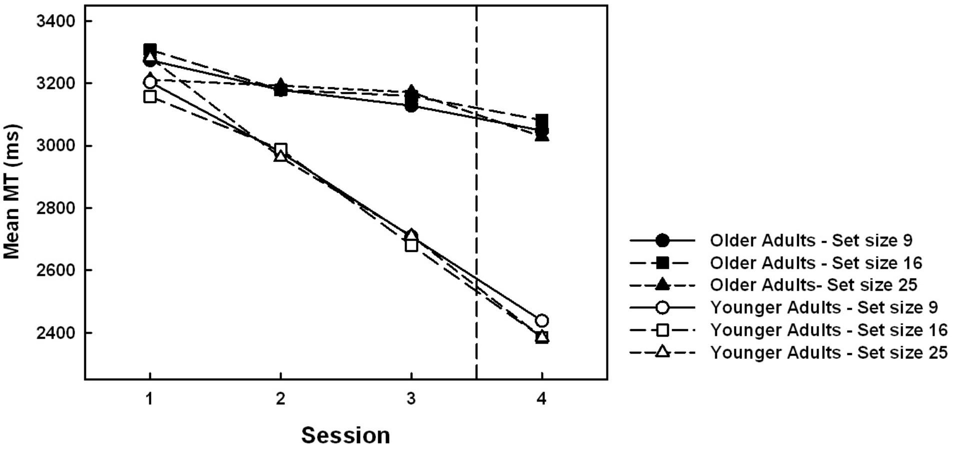


Figure 5.10: Mean Movement Time (MT-LC) as a function of session (1-4), age group (younger, older) and set size (9, 16, 25).

A first indication that aging affects learning in dual-task driving situations comes from driving performance measures. There was a significant main effect of age on MDEV ($F$[1, 18] = 8.46, $\eta_p^2 = .32$, $p < .01$) with MDEV values for older adults (M = 1.27 m) being higher than those of younger adults (M = 1.11 m). Lane-change duration provides important information concerning the effect of aging on practice. Figure 5.10 presents MT-LC as a function of session, age group and set size. The 2-way interaction between set size and age ($F$[3, 51] = 3.20, $\eta_p^2 = .16$, $p = .07$) just missed to reach significance. This close-to-significant interaction was however modulated by a significant 3-way interaction between the factors session, set size and age ($F$[6, 102] = 2.64, $\eta_p^2 = .13$, $p < .05$). To follow up on this interaction two separate 2-way (session x set size) ANOVAs were conducted for each age group respectively. There was no significant interaction between session and set size for older adults ($p > .38$) and this same interaction only showed a tendency for younger adults ($F$[6, 48] = 2.36, $\eta_p^2 = .23$, $p = .09$). The interaction can thus be explained by the fact that practice had no effect on movement times of older adults, independent of the set size presented. Practice tended to have a beneficial effect on younger adults though and although not statistically significant, the benefit seems to be strongest for the larger set sizes.

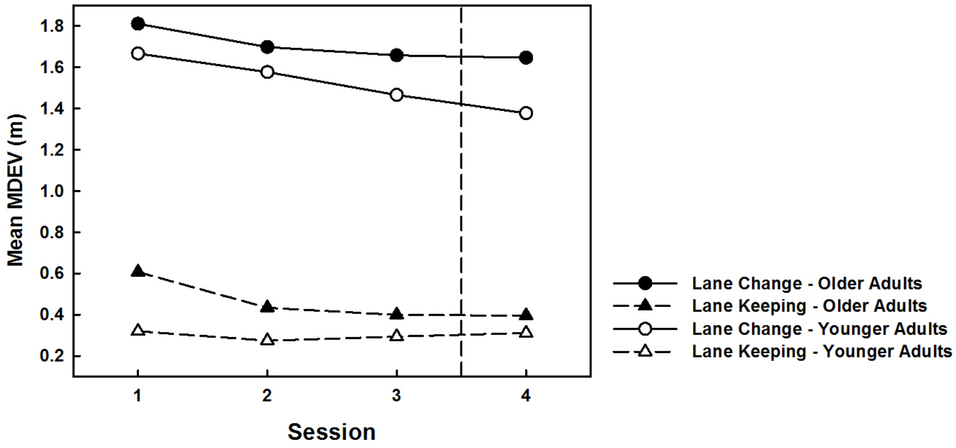


Figure 5.11: Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older).

Figure 5.11 presents mean MDEV as a function of session, segment and age group. The 3-way ANOVA including those factors showed a tendency, but missed to reach significance as well ($F$[3, 54] = 3.01 $\eta_p^2$ = .14, $p$ = .06). It seems to indicate however that over sessions MDEV performance for all road segments and all age groups improved (i.e., lane-change accuracy improved). However, younger adults' MDEV improved more in the LC segments, whereas the older adults benefited more from practice in the LK segments.

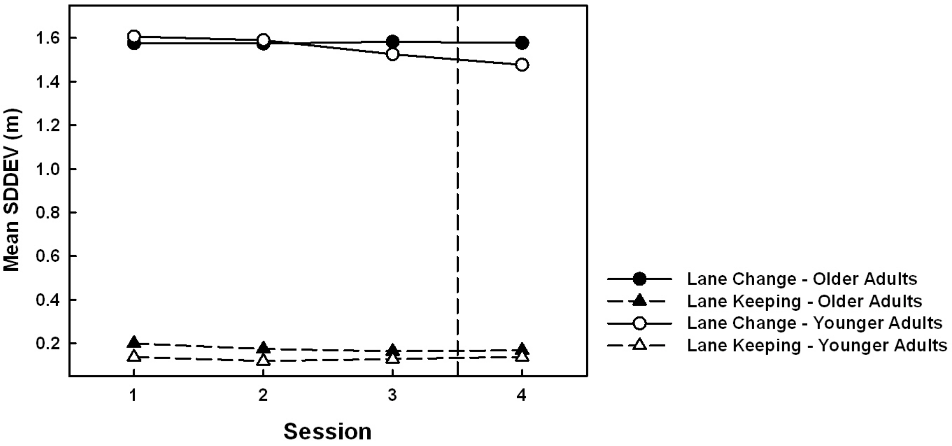


Figure 5.12: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older).

Inherent to MDEV, some evidence that age-related changes affect learning in dual-task driving situations, comes from SDDEV values. Figure 5.12 presents SDDEV as a function of session, segment and age. Although difficult to observe in the figure, there was a significant main effect of session ($F$[3, 54] = 7.60, $\eta_p^2$ = .30, $p < .01$): SDDEV values decreased with practice. The interaction between session and age was significant ($F$[1, 18] = 3.41, $\eta_p^2$ = .16, $p < .05$), but modulated by a significant 3-way interaction including the factors session, segment and age ($F$[3, 54] = 4.34, $\eta_p^2$ = .19, $p < .05$). Follow-up post-hoc tests with two separate 2-way (session x age) ANOVAs for LC and LK segments respectively, revealed that there was no significant interaction between those factors for LK segments ($p > .11$), whereas for LC segments this interaction was significant ($F$[3, 54] = 4.23, $\eta_p^2$ = .19, $p < .05$). Separate one-way (session) ANOVAs for older and younger adults respectively, revealed that there was a simple main effect of session for younger adults ($F[3, 27] = 7.38$, $\eta_p^2 = .45$, $p < .01$), but not for older adults ($p > .34$). The interaction thus shows that only younger adults' performance improves with practice in the LC segments. This is in accordance with the previous findings, and might again confirm that older adults are potentialy "overwhelmed" with the more difficult LC segments, so that little to no learning takes place. An in-depth analysis of this improvement with paired-samples *t*-tests revealed that learning took place between the first and the third session ($t[9] = 2.71$, $p < .05$), but that SDDEV values were the same between the third and fourth session ($p > .25$). In other words, the retention interval had no effect on learning, neither did it affect any previously acquired skills.

Looking at secondary task measures, there was a main effect of age on RT-VST ($F[1, 17] = 55.90$, $\eta_p^2 = .77$, $p < .001$). Older adults needed more time to respond to a visual-search task (M = 1302 ms) as compared to younger adults (M = 924 ms). There was a main effect of age on the proportion of incorrect responses as well ($F[1, 17] = 47.73$, $\eta_p^2 = .74$, $p < .001$): Older adults produced a higher proportion of incorrect responses (M = .25) as compared to younger adults (M = .07).

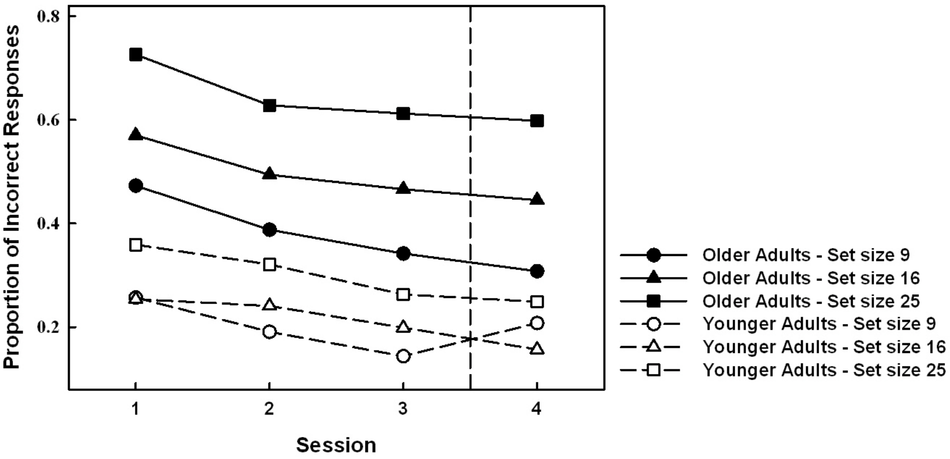


Figure 5.13: Proportion of incorrect responses as a function of session (1-4), age group (young, old) and set size (9, 16, 25).

Figure 5.13 presents the proportion of incorrect responses as a function of session, age group and set size. The 3-way interaction between those factors showed a tendency, but missed to reach significance ($F$[6, 102] = 2.23, $\eta_p^2$ = .12, $p$ = .07) as well. Looking at the figure, it might however provide an indication that age-related changes have an effect on learning in dual-task situations: Older adults produce more incorrect responses than younger adults and their performance seems to improve gradually with practice over sessions. Although younger adults produce much less incorrect responses, their performance improves less over sessions. It should be noted though, that their proportion of incorrect responses is much lower than the proportion of incorrect responses for older adults and therefore maybe does not leave a lot of room for improvement.

Figure 5.5 presents the number of misses as a function of session, segment and age group. The ANOVA revealed a main effect of age ($F$[1, 17] = 11.17, $\eta_p^2$ = .40, $p$ < .01): The proportion of misses produced by older adults was significantly higher (M = .05) than those produced by younger adults (M = .00). The 3-way interaction including all factors (session x segment x age) was significant. A follow-up analysis of this interaction showed that the proportion of misses did not reduce for younger adults with practice (in both LC and LK segments). However, older adults benefited from practice for both segments.

Figure 5.6 presents the proportion of misses as a function of session, age group and set size. The factor set size interacted significantly with the factor age ($F$[2, 34] = 10.45, $\eta_p^2$ = .38, $p$ < .01). Two separate 1-way (set size) ANOVAs for each age group showed that there was a significant simple main effect of set size for older adults ($F$[2, 18] = 13.52, $\eta_p^2$ = .60, $p$ < .01) but not for younger adults ($p$ > .64). An increase in set size thus increased the proportion of misses for older adults, whereas for younger adults an increase in set size had no effect. The 3-way interaction including the factors session, set size and age showed a tendency, but just missed to reach significance ($F$[6, 102] = 2.28, $\eta_p^2$ = .12, $p$ = .06). It seems to indicate however that practice had hardly any effect on the proportion of misses for younger adults, whereas for older adults practice decreased the proportion of misses, especially for larger set sizes. This finding might be due to the fact that younger adults hardly make any misses, and therefore have little room for improvement, especially compared to older adults.

Table 5.7: Results of 3-way (session x task type x age) mixed-factors ANOVAs performed on each of the NASA-TLX subscale ratings. Presented are each of the NASA-TLX subscales, the factors that reached significance or close-to-significance on each of those scales, the degrees of freedom (df), $F$-values, effect sizes ($\eta_p^2$) and probability levels for statistical significance ($p$).

| NASA-TLX Subscale | Factor | df | $F$ | $\eta_p^2$ | $p$ |
|---|---|---|---|---|---|
| Physical Demand | Age | 1,18 | 12.29 | .41 | $< .01$* |
| Temporal Demand | Age | 1,18 | 3.61 | .17 | $= .07$ |
| Performance | Age | 1,18 | 7.16 | .28 | $< .05$* |

Taking into account subjective meausures, some indications for age-related differences in learning can be found. Table 5.7 presents the main effect of age as a result of the 3-way mixed-factors ANOVA on some of the 6 subscales of the NASA-TLX. The factor age had a significant effect on ratings of physical demand, temporal demand and performance. Older adults rated physical and temporal demand significantly higher than younger adults. On the subscale performance, younger adults generated higher values, which for this particular subscale, means that they rated their own performance more severely than older adults did.

Dual-task data, such as IRI as well as eye-movement measures provide indications for the effect of age on learning as well. There was a significant main effect of age on IRI ($F$[1, 17] = 8.13, $\eta_p^2 = .32$, $p < .05$): IRI of older adults was higher (M = 1219 ms) than that of younger adults (M = 921 ms), potentially indicating that they "suffered" more from a secondary task while driving than younger adults.

As to eye-movement data, Figure 5.8 presents VSF as a function of session, age and segment. The ANOVA yielded a main effect of age ($F$[1, 18] = 14.194, $\eta_p^2 = .44$, $p < .01$) with older adults looking down at the search display less often (M = 18) than younger adults (M = 24). As mentioned before, the interaction between session and age was significant. A follow-up analysis showed that the factor session had no effect on younger adults, indicating that the frequency of their vertical saccades did not change with practice. There was however a significant simple main effect of session for older adults, indicating that with practice, VSF for older adults increased. This might reflect that with practice, older adults feel more confident while performing the task and take their eyes more off the road to look at the display more often.

Other eye data indicating a difference in learning between younger and older adults comes from mean glance durations and total glance durations. There was a main effect of age on MGD ($F[1,\, 18] = 38.38$, $\eta_p^2 = .68$, $p < .001$) with MGD of older adults (M = 1041 ms) being significantly longer than those of younger adults (M = 821 ms).

Figure 5.9 presents mean TGD as a function of session and age group. The interaction between both factors was significant. A follow-up analysis showed that the TGD of older adults increased significantly with practice, whereas the TGD of younger adults did not change with practice. Looking at Figure 5.9, one explanation could be that younger adults' total glance duration is already at maximum, leaving no room for improvement.

## 5.4 Discussion

As to the first research question, whether learning had an effect on dual-task driving performance, this experiment showed that indeed, learning has an effect. Generally, this effect was positive, in that performance improved with repeated practice. Learning effects could be observed in all driving performance measures: Lane-change initiation times, movement times, lane-change accuracy and steering-wheel variability improved. Learning had a positive effect on secondary task performance measures as well: Reaction times and both the number of incorrect responses and misses decreased with practice. Finally, dual-task data, such as the Inter-Response Interval (IRI), a variable dedicated to measuring dual-task interference, witnessed of the effect of practice on performance. Analyses revealed that with practice, IRI values and thus dual-task interference, decreased. These objective findings are backed up by subjective findings: Participants progressively decreased their ratings of mental effort (mental demand, temporal demand), and they subjectively rated the tasks as demanding less effort and frustration. These findings are in accordance with Anderson's ACT-R theory (1982) as well as Rasmussen's SKR-Model (Rasmussen, 1983), according to which tasks become (partially) automated with practice, thereby freeing up resources, leading to improvements on all performance measures.

It should be noted though, that those learning effects were often modulated by other factors such as driving task difficulty (LC versus LK segments), the addition of a secondary task as well as age effects (the latter one will be referred to in the sections below). As such, lane-change accuracy (expressed by MDEV) was influenced by driving task complexity (i.e., the more difficult lane-change segments versus the easier lane-keeping segments): Although learning took place in both road segments, the strongest learning effects were found in LC segments. These findings are as expected: As LC segments are more difficult, at first those segments provoke worst performance outcomes than LK segments. With practice however, there is ample room for improvement, leading to stronger learning effects as compared to LK segments, where performance is better from the first session on. Lane-change accuracy (MDEV) as well as lateral deviation variability (i.e., stability as expressed by SDDEV) were both influenced by task-type difficulty as well: Stronger learning effects were observed in the single-LCT condition as compared to the dual-task condition. As to the differences in learning rates of driving performance measures between the single-LCT condition and the dual-task condition, it seems that the addition of a secondary task represented such an important increase in task complexity, that even when instructed to focus on the primary task, some performance decline was observed, leading to slower learning rates for driving performance measures.

A more in-depth analysis of learning effects showed that learning took place between the first and the third session, most of the time independent of related factors such as task type difficulty (single versus dual) or driving task complexity (lane change versus lane keeping). The retention interval however had no effect on any of the performance measures: Acquired skills did not deteriorate with a retention interval, neither did they improve. This is an indication that acquired skills do not diminish over time, once they have been integrated.

As to the second research question, whether age-related changes affect learning in dual-task driving situations, this experiment showed as well that differences in learning can be found between younger and older adults. As mentioned in the "General Effects" section (Chapter 5.3.1), age had an effect on almost all performance measures, leading to younger adults performing in general better than older adults. As to the effects of aging on learning in dual-task driving environments, some differences between age groups were observed. As such, lane-change accuracy (as expressed by MDEV) improved most for younger adults in LC segments, whereas the biggest improvement in MDEV values for older adults was observed in LK segments. Similar results could be found for SDDEV values: Lateral deviation variability decreased for younger adults in LC segments, but showed no learning effect for older adults. In other words, younger adults benefit from practice not only by improving their lane-change accuracy performance (MDEV) but also by becoming less variable (i.e., more stable) in their lane-change behavior (SDDEV), and this especially in the more difficult LC segments. Older adults' learning, and therewith improved performance on driving measures, can especially be observed in the easier LK segments. One explanation why no learning effects for younger adults on those performance measures were observed in LK segments, is that those values were already very low and left little potential for improvement (i.e., a ceiling effect). As to the question why older adults do not seem to benefit from practice in the more difficult LC segments, one explanation could be that the driving task was so complex for them, that even with practice, they could not free any resources, leading to a stagnation in dual-task driving performance. Another explanation could be that older adults do not manage to integrate the SKR feedback in a manner as to improve their performance like younger adults do.

Another strong effect of age on learning was found in the proportion of misses on the VST produced. With practice, the proportion of misses on the VST decreased for older adults, whereas no effect of practice on the proportion of misses was found for younger adults. One plausible explanation for this latter finding is the fact that younger adults hardly produced any misses and that therefore their performance on this measure did not improve (a lot) with practice. Older adults however, produced an important number of misses, especially in LC segments, but in LK segments as well. One explanation for these findings might come from the task instructions used: Participants were instructed to prioritize the driving task and even ignore the secondary task if they thought they could not handle the situation. Especially in LC segments, older adults might have been overwhelmed with the dual-task situation, resulting in ignorance of the visual search task. This might have created a high number of misses in the first session, creating a bias in the data set. The decrease in the proportion of misses was strongest for bigger set sizes. A plausible explanation for these findings is that a ceiling effect occurs from the first session on (i.e., the smaller set sizes did not produce a high amount of misses), therewith leaving little to no room for measurable improvement.

It thus seems that, despite practice and despite improvements on some of the performance measures, older adults do not benefit as much from practice as younger adults do and their performance always lags behind performance of younger adults. This is in accordance with findings by other authors (for example Caird et al., 2008; Göthe et al., 2007; Verwey et al., 2011). One reason for the age differences observed in skill acquisition may be the age-related slowing in cognitive processes (Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) such as attentional processes and memory functions, which are important for skill acquisition (Anderson, 1982). Another explanation for performance differences between younger and older adults might come from eye-movement data. With practice, vertical saccade frequencies do not change for younger adults, whereas for older adults they increase over sessions. Total glance duration of older adults increased significantly with practice as well, whereas the total glance of younger adults did not change. One reason for the repeated absence of learning effects in eye data for younger adults, could be that they were already at a plateau of performance, leaving little to no room for improvement. The observed changes for older adults though might potentially indicate a change in dual-task strategy with practice: With practice, parts of the (driving) task get automated, therewith freeing up resources, which allow more efficient time-sharing with secondary tasks. Indeed, older adults might feel more confident (as witnessed by decreases in subjective ratings concerning mental demand and effort with practice) over sessions and turn to a more parallel processing instead. These findings do not lead to any changes in performance outcomes, but indicate that with practice, older adults change strategy to handle dual-task situations.

A finding of interest is the effect, or rather the absence of effect, of the retention period on learning. There was no difference in performance after a retention period of three weeks for both younger and older adults. This might be an indication that the three sessions used in this experiment were sufficient to establish a stable learning outcome (insensitive to decay) for both younger and older adults.

## 5.5 Conclusion

Taken together, it can be said that learning indeed has a (positive) effect on dual-task driving performance: Driving as well as secondary task measures improve with practice over sessions, therewith improving dual-task performance as a whole. A longer period without practice had no negative effect on retention: Participants' performance remained at the same level despite a retention period. In the light of IDSS within a vehicle, those findings are positive, as they indicate that the use of IDSS in parallel to the driving task can be learned and that moreover, acquired skills are not immediately forgotten once practice stops.

Learning was age-dependent though: Younger adults learned at a faster rate than older adults and their performance on all tasks was always a bit better than that of older adults as well. Older adults furthermore benefited less from practice in more difficult situations (e.g., in more difficult driving situations or general increases in task complexity by adding a secondary task) as compared to younger adults. These findings are partially in accordance with previous studies examining differences in dual-task performance (combining a sensory-motor task with a cognitive task) between age groups (Voelcker-Rehage & Alberts, 2007), showing that age-related cognitive changes affect skill acquisition, due to restraints in cognitive functions (Repovs & Baddeley, 2006) and a lack of transition from serial to parallel processing (Göthe et al., 2007). However, this experiment seems to provide first indications, based on eye-movement behavior that with practice, older adults make a transition from serial to parallel processing. This change in strategy furthermore does not seem to be affected by a retention interval, providing an indication that changes are rather permanent.

One limitation of this study, is that the secondary task used was of no relevance for the driving task. In other words, participants could ignore the secondary task completely and still be able to correctly perform the driving task. In real-world conditions, this is rarely the case: IDSS are implemented to provide information or assistance and are (generally) not supposed to be ignored. One interesting research question that remains is therefore whether secondary-task relevance influences learning performance in a dual-task driving situation. In the experiment that follows, this research question will be considered more into detail.

# 6 Experiment 4: Aging and Learning in Relevant Dual-Task Driving Situations

## 6.1 Introduction

The goal of this experiment was to examine the effect of relevant secondary tasks on the driving task. In real driving conditions, an IDSS has the goal to provide information or provide assistance. Ignorance of the IDSS by the driver might lead to discomfort (e.g., when navigation instructions are ignored) or even to dangerous situations (e.g., when warnings of a Lane Change Assistant are ignored). In the previous experiment, the surrogate IDSS, in the form of a visual search task was of no relevance for the driving task: In difficult situations, participants could choose to even ignore it without subsequent consequences for the driving task. The goal of this experiment is to examine whether driving with a relevant secondary task has an effect on learning in dual-task driving situations for younger and older adults. More precisely, the following research questions were examined: Does secondary-task relevance influence performance in a dual-task driving situation? Does dual-task performance differ between younger and older adults with a relevant secondary task? Does the use of a relevant secondary task have an effect on learning in dual-task driving situations? And finally, does the use of a relevant secondary task have an effect on learning between age groups?

A meta-analysis by Seppelt and Wickens (2003) reviewing studies using relevant secondary tasks or messages (i.e., tasks or messages delivering information which is "related" to the primary driving task, such as navigation, road and traffic information) revealed that relevant visual secondary tasks might be more distracting than irrelevant (i.e., messages and information that refer to infotainment messages) visual secondary tasks or messages. For irrelevant information, the driver needs to be able to ignore the secondary task, which, with a visual message is easier, as the driver can turn his gaze away from the secondary task display and onto the road. However, with a relevant secondary message or task, not only the driver needs to direct his attention off the road and towards the message or task, but visual messages might also be ignored if the driving task gets too demanding (note that for this reason, it is recommended to present safety-relevant messages in an auditory manner; Wickens & Seppelt, 2002).

Seppelt and Wickens (2003) furthermore conducted a study themselves in which it was revealed that secondary task relevance affected driving performance measures in a high-fidelity driving simulator. The secondary task consisted of visual and auditory messages that carried either direct relevance to the driving task (i.e., warning messages of driving events ahead like for example "Accident in road ahead.") or that carried no direct relevance to the driving task (e.g., "Mostly sunny and warm. Highs in the 70s."). To assure participants had read or heard both relevant and irrelevant messages, they were asked to repeat them out-loud. As expected, based on the multiple-resource theory (Wickens, 1984) as well as the findings of previous studies (Seppelt & Wickens, 2003), the benefits of auditory delivery of information were especially strong when that information was relevant for the driving task as compared to when that information was irrelevant. The researchers found however an interesting main effect of message relevance on steering-wheel velocity as well: Relevant messages produced less steering-wheel velocity than irrelevant messages. The authors interpret this as an indication that drivers had less aggressive steering control when attempting to process a relevant message than when processing an irrelevant message. The authors explain these findings by the fact that relevant messages kept the driver's attention more focused on the roadway and hence caused less distraction than an irrelevant message. The authors found as well indications that drivers were able to "disengage" from an in-vehicle task when it was irrelevant and the driving task was more demanding (i.e., with a hazard up ahead).

To our knowledge, no study has examined how aging or learning affects the effect of secondary task relevance on driving performance. To examine these questions more into detail, an experiment was conducted in which a group of younger and older participants had to perform the LCT (Mattes, 2003) in combination with a visual search task (Treisman & Gelade, 1980), which was either relevant or irrelevant for the driving task.

### 6.1.1 Research Questions and Hypotheses

The first research question examines whether a relevant secondary task influences performance in dual-task driving situations. In accordance with findings by Seppelt and Wickens (2003) and based on the multiple-resource theory (Wickens, 1984), it is expected that a relevant visual search task will interfere more with the visual driving task than an irrelevant visual search task. Visual attention will be forced off the road to attend to the visual search task, which is expected to have a negative effect on driving performance measures. As the relevant secondary task cannot be ignored, it is also expected that reaction times on the relevant visual search task will increase as compared to an irrelevant secondary task. Based on the nature of the secondary task (i.e., a visual search task which participants have to attend to), interference effects on the relevant secondary task were expected to be even stronger than effects found by Seppelt and Wickens (2003), who used messages which could be ignored, without affecting the driving task (even the relevant messages). Based on these differences, the visual secondary task was expected to have a larger impact on the dual-task driving condition than findings by those authors, leading to even stronger negative impacts on driving performance measures as well as secondary task measures.

As to the second research question, whether performance differs between younger and older adults with a relevant secondary task, age effects are expected. A visual search task demands discerning a target in between distractor items, for which it is known from previous studies that older adults need more time (Hahn et al., 2011). Indeed, the previous study in this thesis showed that older adults need more time to solve a visual search task. If the visual search task now becomes of relevance for the driving task (i.e., by providing information for the next lane-change maneuver) differences between older and younger adults might become more obvious, as compared to situations in which the secondary task is irrelevant for the driving task. Especially differences in driving performance measures are expected: As more attentional resources need to be directed onto the secondary task, interference on the primary task will increase. In accordance with findings in Experiment 3, this interference is expected to be higher for older adults, as they suffer more from the addition of a secondary task than younger adults do.

The third research question examines whether the use of a relevant secondary task has an effect on learning in dual-task driving situations. Based on the previous experiment in this document, performance on both the driving as well as the secondary task was expected to improve with practice. A theoretical foundation for this hypothesis comes from the ACT-R theory by Anderson (1982) according to which tasks become (partially) automated with practice, thereby freeing up resources. However, if the relevant secondary task really distracts more than the irrelevant secondary task, it can be expected that learning curves for both driving as well as secondary task performances will be less steep than learning curves for settings in which the secondary task is potentially less distracting (i.e., irrelevant for the driving task).

The final research question examines whether the use of a driving-task relevant secondary task has an effect on learning between age groups. As mentioned before, some learning is expected, as practice allows for automatization (Anderson, 1982), leading eventually to better time-sharing between tasks. Learning rates are expected to differ though between younger and older adults, due to age-related changes in cognitive processes needed for learning (e.g., attentional processes, memory functions; Anderson, 1982), for solving visual search tasks (e.g., working memory functions needed to discern a target from distractor items; Hahn et al., 2011) and managing dual-task situations (Wilschut et al., 2008). Based on these theoretical assumptions as well as results in Experiment 3, it is expected that older adults will learn at a slower rate than younger adults. This difference in learning between age groups will be visible in both primary as well as secondary task performance measures and is expected to be stronger in the condition with a relevant secondary task.

## 6.2 Method

### 6.2.1 Participants

A total of 34 individuals (17 younger participants, mean age = 26.7 years, age range = 24-34 years; 17 older participants, mean age = 66.8 years, age range = 64-75 years) participated in the study. Younger drivers reported driving 11176 km per year and were in possession of a driver license for a mean period of 8.3 years. Older participants reported driving 11765 km per year and their mean period of possession of a driver license was 42.9 years.

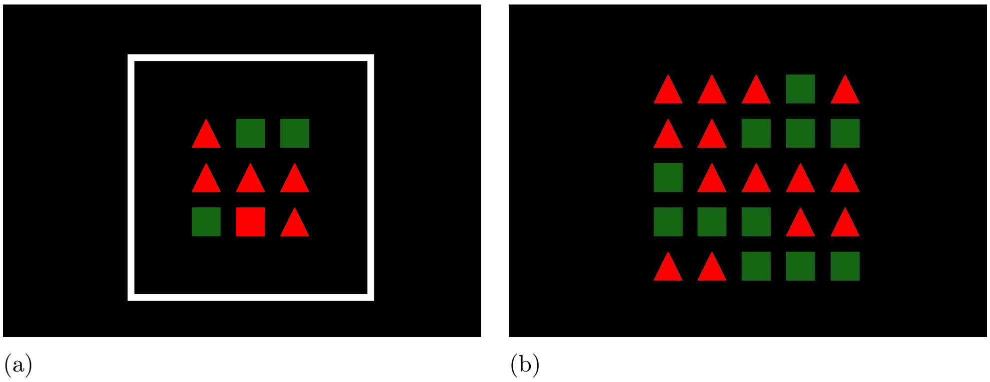
### 6.2.2 Procedure

To account for effects of practice and learning in a situation where the secondary task is of relevance for the primary task, the experiment consisted of 5 sessions. The first session was a practice session which took approximately 1.5 hours. In this session, participants read and signed a consent form, filled out the demographic questionnaire and performed some neuropsychological and vision tests as described in the General Method section (Chapter 2.3). They were then introduced to the different tasks within the experiment and had the possibility to perform up to 3 practice blocks for each task to ensure all tasks were well understood.

The practice session was followed by 4 experimental sessions of approximately 3 hours each. To allow for the registration of eye movements, electrooculogram (EOG) recording electrodes were attached to the participant's face before starting the first task. All sessions took place within a timeframe of 3 weeks with an average inter-session interval of 3 days.

For the purpose of this experiment, the VST was changed into a Go-NoGo task with either 9 or 25 items. The search display with 16 items was discarded for this experiment. When a target was present on the search display, participants had to push either of the two buttons on the back of the steering-wheel and they would hear a short beep (400 ms) to confirm button press. If no target was present, no button was to be pressed and no confirmation beep would sound.

Within one session, the single-session experimental design was applied as described in the General Method section (Chapter 2.3) for the single-LCT and the single-VST condition. As such, participants performed 5 single-LCT experimental blocks, 5 single-VST experimental blocks and 10 dual-task experimental blocks (this setting will be referred to as "dual"). They then performed 10 dual-task experimental blocks in which the VST became relevant for the primary driving task. This setting will from now on be referred to as ”dual-r”. Just as in the dual setting, in the dual-r setting participants were instructed to look for a target on a visual search display and push a button accordingly,



(a) (b)

Figure 6.1: Examples of an (a) relevant the Visual Search Task (VST) and b non-relevant VST.

as described just above. VSTs were presented both when lane changing (LC-VST) as well as when lane keeping (LK-VST). However, only the participant's reaction on the LC-VST was of relevance for the next lane change: If a button had been pressed (and thus a confirmation beep had sounded) on the LC-VST, the participant had to follow the instructions on the next lane-change sign. If no button had been pressed (and thus no beep had been produced) on the LC-VST, the instructions on the next lane-change sign were to be ignored and the participant should continue straight on his current lane. The LK-VST needed to be performed as fast and as accurate as possible as well, but was of no relevance for the driving task; it served as a distracter (interference) task only. To help participants remember which was the relevant secondary task (i.e., the task that had a direct impact on the driving task), a white border appeared around the LC-VST search task (see Figure 6.1).

All participants received end-of-block feedback over their driving and secondary-task performance as described in Experiment 2: Effect of Feedback on Dual-Task Performance (Chapter 4).

Taking into account the changes described in the previous paragraph, the experiment consisted of a mixed-factors design including 1 between-participants factor and 5 within-participants factors. First the between-participants factor *age* to reflect the effect of age on learning effects. Within-participants factors consisted of the number of sessions (*session*, 1-4) to examine possible learning effects by practice, the factor *task type* to reflect task-type differences (single-LCT, single-VST, dual, dual-r), the factor *segment* (lane change, lane keeping) to examine differences in driving-task complexity and finally the factor *set size* (9, 25) to reflect differences in secondary task complexity.

### 6.2.3 Data Analysis

All experimental data from the first (practice) session were discarded for analysis. As the experimental setting differed from previous settings, LCT-data were discarded as follows. First each driving trial was categorized into one of four categories:

- "Correct - Change" when the participant had to change of lane (in any direction and over 1 or 2 lanes) and F changed lane correctly.
- "Correct - No Change" (CNC) when the participant had to stay in his lane and indeed stayed in his lane.
- "LCT-Error" when the participant had to change of lane, but either changed into the wrong direction or the wrong number of lanes.
- "LCT-Miss" when the participant was supposed to stay on his lane (i.e., drive straight on), but changed lane instead.

All trials "Correct - No Change" were discarded from analysis as they did not provide any useful data. As can be seen in Table 6.1 below, the amount of "Correct - No Change" trials discarded are quite high, as they theoretically correspond to at least half of the trials. Then all other trials including the driving task were excluded following different exclusion criteria (including LCT-Error and LCT-Miss) described in the Data Analysis section of the General Method (Chapter 2.4). As LK sections were not analyzed in this experiment, the filter "TLKT" was not applied. Table 6.1 presents the percentage of excluded trials per age group ($N_{\text{Old}} = 17$, $N_{\text{Young}} = 17$) and trial type (LCT, DUAL, DUAL-R).

Table 6.1: Percentage of excluded trials for the driving (LCT) data, following six exclusion criteria described above and in the Data Analysis section (Chapter 2.4) of the General Method.

| Age Group | Con | CNC | AE | RT-LC | MT-LC | LCT-Error | LCT-Miss |
|---|---|---|---|---|---|---|---|
| Young | LCT | - | 0.06 | 0.15 | 0.04 | 0.02 | .00 |
| Young | DUAL | - | 0.12 | 0.58 | 0.25 | 0.36 | .00 |
| Young | DUAL-R | 56.46 | 0.33 | 2.78 | 0.35 | 3.58 | .00 |
| Old | LCT | - | 0.13 | 2.56 | 0.25 | 0.13 | .00 |
| Old | DUAL | - | 0.74 | 4.28 | 1.01 | 1.85 | .00 |
| Old | DUAL-R | 60.62 | 0.92 | 6.16 | 0.68 | 5.03 | .00 |

*Note.* Con = Condition; CNC = Correct - No Change; AE = Algorithm Error; RT-LC = Reation-Time until Lane Change; MT-LC = Movement Time until Lane Change.

For the VST data, only trials in which a correct response was given and where a target had been present were considered for analysis (i.e., half of the trials). From the remaining trials, all trials in which the participant had answered faster than 200 ms were discarded as they were considered anticipations. In the single-VST condition, 2.78% of the remaining trials were discarded for younger adults and 15.54% for older adults. In the dual condition, 4.65% of the trials were discarded due to anticipation for younger adults and 20.66% of the trials for older adults. In the dual-r condition, 3.79% of the trials were discarded due to anticipation for younger adults and 22.68% for the older adults.

## 6.3 Results

In this results section, due to the large amount of data, results are presented as follows: First general effects are reported, which are expected based on previous research, but which show that this study is consistent with previous studies ("General Effects"). General effects will only be reported, without providing any statistics or figures (the according inferential statistics are provided in Appendix D).Then, only results of relevance for answering the research questions are presented. An overview of the complete data analysis (including statistical analyses and figures for general effects), can be found in

Appendix D. Note that due to the amount of data, figures are only included for significant or close to significant ($p = .08$) main effects or interactions. Close to significant interactions will only be described, without providing any exploratory follow-up tests. If an interaction accounts for answering two or more research questions, the statistical analysis of this interaction will only be done once, when the interaction is encountered for the first time. Later references to that same interaction will simply be referred to and main findings will be resumed, before being interpreted for the research question at hand. If a significant main effect or a significant simple main effect with the factor *session* was present, power functions were fitted to illustrate how performance changes with learning. The equation used for fitting was the following:

$$y = a + bx^{-c}$$

where $y$ represents a dependent variable, $x$ represents the session number, $a$ represents the asymptote, $b$ is the amount of time required at the beginning of training and $c$ specifies the rate of speed-up with training. $R^2$ and Root Mean Square Error (RMSE) were used to quantify goodness of fit. Arcsine-transformed data for which power functions needed to be fitted, are always represented with two graphs: One representing the non-arcsine transformed data (as this is what is typically presented) and another one representing the arcsine-transformed data including the fitted curves (as these are the actual data the ANOVAs were performed on).

## 6.3.1 General Effects

Due to the amount of data, general effects are divided into findings concerning general task complexity, age effects, secondary task complexity and learning effects.

### General Task Complexity

As expected, there was an effect of general task complexity (i.e., single- versus dual-task data) on driving performance measures, secondary task measures and subjective measures. Onset times until lane change (RT-LC), lane-change movement times (MT-LC) and lane-change accuracy (as witnessed by MDEV and SDDEV) were highest in the dual-r condition, followed by the dual condition and finally the single-LCT condition. Only MT-LC findings were influenced by age effects: With increasing task-type difficulty (i.e., single versus dual-task settings), the MT-LC values for older adults increased more than those of younger adults. Data on steering-wheel variability showed similar resultss: Steering-wheel variability was lowest in the dual-r condition, increased in the dual condition and was highest in the single-LCT condition. This is in accordance with previous findings by Berthon-Donk and colleagues (2011) as well as the previous experiment in this thesis. According to the abovementioned authors, steering-wheel variability represents the number of corrective steering-wheel corrections. These are known to be higher in easier (driving) conditions.

Secondary task performance measures reflected the influence of task complexity on performance as well: Reaction times on the visual search task were highest in the dual-r condition, followed by the dual condition and finally the single-task condition. The proportion of incorrect responses showed a similar pattern for older adults: With an increase in overall task-difficulty, the proportion of incorrect responses increased. An increase in task difficulty had statistically no effect on the proportion of errors produced by younger adults though. The proportion of misses however, increased for both age groups with an increase in task-complexity. This increase was stronger for both age groups with the bigger set size as compared to the smaller set size.

An analysis of subjective measures confirmed the above main effects of task type difficulty: Participants rated the single-LCT condition as least demanding, followed by the single-VST condition, the dual and finally the dual-r condition. Similar results were found for the NASA-TLX on the sub-scales physical demand, temporal demand, performance and frustration. Surprisingly, participants rated the single-LCT as well as the single-VST condition as demanding more effort than the dual and the dual-r conditions. One explanation for these rather unexpected findings could be the blocked order of the experiment, leading to a form of primacy effect (Ebbinghaus, 1885b), leading to higher ratings for the first conditions as compared to later conditions. Participants first performed the single-LCT and the single-VST conditions, before being introduced to the more difficult dual-task conditions (dual and dual-r). For this reason, they might have "overrated" their effort at first, leading to lower values for the more difficult conditions later on. Ratings on the NASA-TLX scale mental demand, differed between age groups: With increasing overall task difficulty, subjective ratings increased more for younger adults as compared to older adults. This might be an indication that older adults overestimate their own capacities. (Kruger & Dunning, 1999).

### Age Effects

As expected, and in accordance with findings in the previous experiments, general age effects were found and reflected in most of the performance measures. As such, older adults were slower to initiate a lane change (as witnessed by RT-LC values) and their lane changes took in general more time (as witnessed by MT-LC values). Their lane-change accuracy was reduced as well, as compared to younger adults (as witnessed by higher MDEV and SDDEV values). As expected, standard deviation of steering-wheel variability (SDSW) showed age effects in accordance with above findings: SDSW values were higher for younger adults as compared to younger adults cite <which is in accordance with findings by>Berthon-Donk2011. Older adults also produced significantly more LCT-Errors and LCT-Misses as compared to younger adults. On secondary task measures age effects were reflected as follows: Reaction times, incorrect responses and misses on the visual search task were higher for older adults as compared to younger adults. Subjective ratings of mental effort were the only variables which did not reflect any general effect of the factor age: Subjective ratings of mental effort did not differ statistically between younger and older adults. However, the NASA-TLX scale for effort found that older adults rated their estimated effort as lower than younger adults. This might be an indication that older adults overrate their own capacities (an effect known as the Dunning-Kruger effect; Kruger & Dunning, 1999).

### Secondary Task Complexity

An increase in secondary task complexity, represented by an increase in set size, had an effect on primary performance as well. These findings are as expected and in accordance with previous experiments in this thesis as well as findings by other researchers (Merat et al., 2005; Wilschut et al., 2008). IRI, an indicator for interference increases with increased set size. Increases in set size especially affected older adults: The proportion of LCT-Errors increased significantly for older adults, but not for younger adults. This might be an indication that older adults "suffer" more from an increase in secondary task complexity than younger adults.

### Learning Effects

Finally, general learning effects were found in this experiment. Older adults especially seemed to benefit from practice, as witnessed by improved driving (RT-LC, MT-LC, MDEV and SDDEV) and secondary task (RT-VST, errors and misses on the VST) performance measures. One explanation for the differences in performance improvement for younger adults is that they are already on a plateau on the driving task, leaving them with ample capacities to concentrate on secondary task performance, hence leading to performance improvements on that task. Older adults not only need to concentrate on improving their secondary task performance, but their driving task performance as well, hence demanding more cognitive resources. As the driving task is more important (experimental instructions!), they focus foremost on improving the primary task and only then on the secondary task. This might lead to improvements in performance at the secondary task at a lower rate.

The observed learning effects were often influenced by an effect of set size and/or age: As such, data showed that reaction times on the visual search task, decreased with practice for both younger and older adults and for all set sizes. However, the speed at which learning took place differed between age group and set size: Learning was fastest for the younger adults and the smallest set size. Learning slowed down for the younger adults and the bigger set size. Finally, learning was slowest for the oldest adults for both set sizes. Similar interactions between those factors were found for the proportion of incorrect responses and misses on the VST.

Learning effects were influenced by both task type and age as well: With practice, the proportion of misses on the visual search-task reduced significantly for both younger and older adults and for all task-type conditions. This reduction was strongest though for older adults in both the dual and dual-r conditions. Older adults seem to suffer most from more difficult task conditions. Subjective findings seem to confirm these observations: Although subjective ratings did not change for younger adults (independent of practice or task type), subjective ratings of older adults dropped over sessions. This drop in subjective rating was strongest for the dual-r condition, followed by the dual condition and finally the single-VST condition. Subjective ratings of mental effort on the single-LCT did not change with practice.

### 6.3.2 First Research Question

The first research question examines whether secondary-task relevance influences performance in dual-task driving situations.

Looking into dual-task data into detail (e.g., comparing dual to dual-r data), secondary-task relevance was found to have an effect especially on driving performance measures. A first indication comes from the number of LCT-Errors produced (i.e., when the participant either changed into the wrong direction or the wrong number of lanes). There was a significant main effect of dual-task type on LCT-Errors ($F[1, 32] = 173.89$, $\eta_p^2 = .85$, $p < .001$): The number of LCT-errors were significantly higher in the dual-r condition (M = .13) as compared to the dual condition (M = .01). Similar results were found for LCT-Misses (i.e., when the participant was supposed to stay on his lane, that is drive straight on, but changed of lane instead). Significantly ($F[1, 32] = 99.97$, $\eta_p^2 = .76$, $p <$ .001) more LCT-Misses were produced in the dual-r condition (M = .04) as compared to the dual condition (M = .01). These findings indicate that participants either missed the instructions from the secondary task (e.g., because they had no time to resolve it) or that they had ignored the secondary task, leading to a potential increase in LCT-Errors and/or LCT-Misses.

Lane-change onset times (i.e., RT-LC) were significantly higher ($F[1, 32] = 28.127$, $\eta_p^2$ = .47, $p < .001$) in the dual-r condition (M = 995 ms) as compared to the dual condition (M = 955 ms). Similar findings were found for MDEV as well as SDDEV values: The ANOVA showed a significant main effect of dual-task type on MDEV ($F[1, 32] = 31.22$, $\eta_p^2 = .50$, $p < .001$). Mean MDEV values were significantly higher in the dual-r condition (M = 1.92 m) as compared to the dual condition (M = 1.83 m). There was a significant main effect of dual-task type on SDDEV ($F$[1, 32] = 31.11, $\eta_p^2 = .49$, $p < .001$) as well: Mean SDDEV values were significantly higher in the dual-r condition (M = 1.54 m) as compared to the dual condition (M = 1.46 m). These findings indicate that dual-task type has an effect on driving performance measures. In the dual-r condition in which the secondary task is relevant for the driving task, driving performance measures decreased as compared to the dual condition in which the secondary task was of no relevance for the driving task.

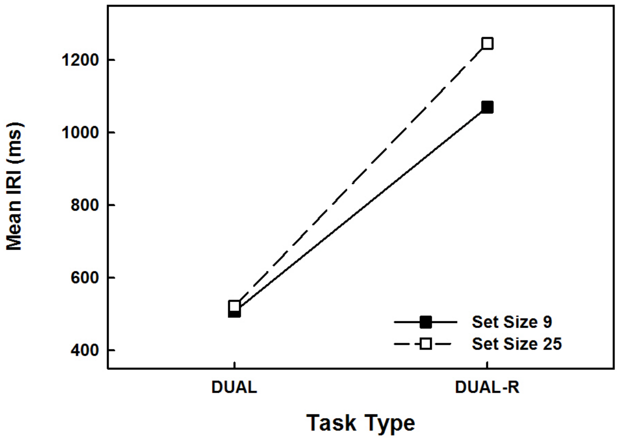


Figure 6.2: Mean Inter-Response Interval (IRI) as a function of dual-task type (dual, dual-r) and set size (9, 25).

Finally, looking at IRI, an indicator for dual-task interference, it turned out that there was an effect of dual-task type on IRI ($F$[1, 32] = 31.11, $\eta_p^2 = .49$, $p < .001$): Mean IRI values were significantly higher in the dual-r condition (M = 1158 ms) as compared to the dual condition (M = 514 ms). This indicates that the relevant secondary task provoked more interference than the non-relevant secondary task. It should be noted though that this finding was influenced by secondary-task difficulty. As such, Figure 6.2 presents mean IRI as a function of dual-task type and set size. The 2-way interaction between those factors was significant ($F$[1, 32] = 36.64, $\eta_p^2 = .53$, $p < .001$). A follow-up analysis with separate 1-way (dual-task type) ANOVAs for set size 9 and 25 respectively, revealed a significant main effect for the visual search display containing 9 items ($F$[1, 33] = 313.33, $\eta_p^2 = .91$, $p < .001$) as well as the visual search display containing 25 items ($F$[1, 33] = 322.11, $\eta_p^2 = .91$, $p < .001$). As can be seen in Figure 6.2, interference increased going from the dual to the dual-r condition. This increase was however strongest when the bigger set size was presented.

### 6.3.3 Second Research Question

The second research question examines whether performance differs between younger and older adults with a driving-task relevant secondary task.

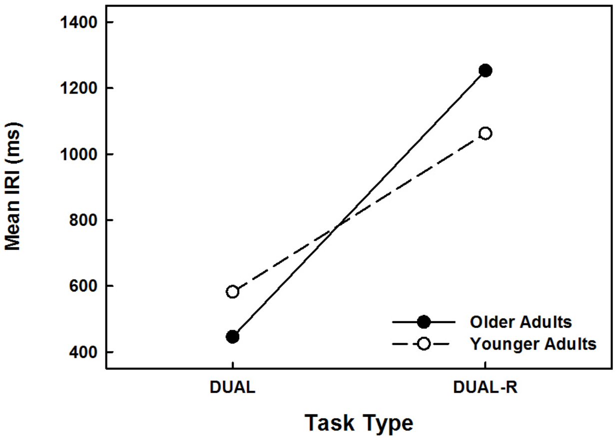


Figure 6.3: Mean Inter-Response Interval (IRI) as a function of (a) dual-task type (dual, dual-r) and age (younger, older).

A first analysis of the dual-task data (i.e., comparing dual to dual-r data) showed that performance differs between younger and older adults when the secondary task is relevant for the driving task. Figure 6.3 presents mean IRI as a function of dual-task type and age. The 2-way interaction between those factors was significant ($F$[1, 32] = 78.86, $\eta_p^2 = .71$, $p < .001$). Follow-up tests with separate 1-way (dual-task type) ANOVAs for older and younger adults respectively, revealed a simple main effect of dual-task type for older adults ($F$[1, 16] = 613.51, $\eta_p^2 = .98$, $p < .001$) as well as for younger adults ($F$[1, 16] = 704.94, $\eta_p^2 = .98$, $p < .001$). Independent samples $t$-tests comparing IRI values between age groups for the dual, as well as the dual-r condition, revealed that for both conditions, differences between age groups were significant (dual: $t[32] = 2.59$, $p < .01$; dual-r: $t[32] = 2.98$, $p < .01$). As can be seen in Figure 6.3, younger adults thus suffer more from the dual as compared to the dual-r situation as opposed to older adults, who suffer more from the dual-r condition as compared to the dual condition. Furthermore, an increase in dual-task complexity has a stronger effect on performance of older adults as compared to younger adults.

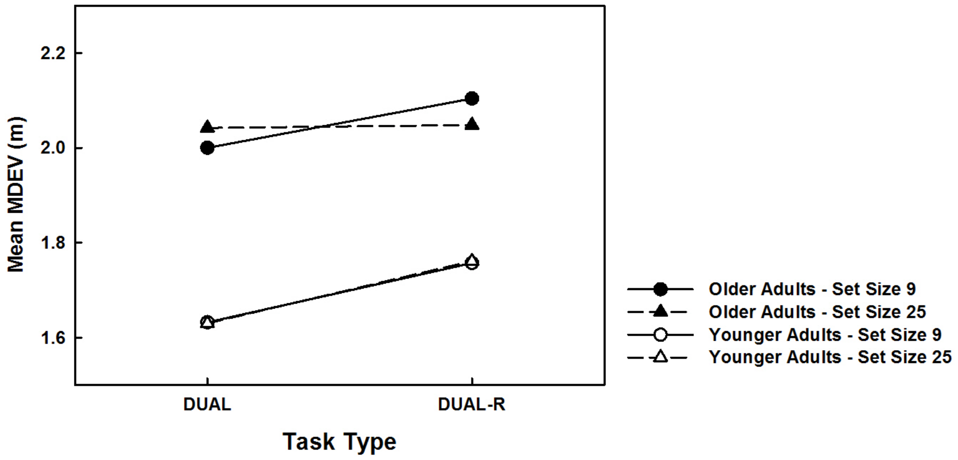


Figure 6.4: Mean Lateral Deviation (MDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25).

Further analysis revealed similar results for lane-change accuracy (as expressed by MDEV and SDDEV), but these interactions were always influenced by set size as well. Figure 6.4 presents mean MDEV as a function of dual-task type, age and set size. The interaction between dual-task type and age was significant ($F[1, 32] = 5.08$, $\eta_p^2 = .14$, $p < .05$). The interaction between dual-task type and set size showed a tendency but missed to reach significance ($p = .06$). Those 2-way interactions were however modulated by a significant 3-way interaction between all factors ($F[1, 32] = 4.69$, $\eta_p^2 = .13$, $p < .05$). A post-hoc analysis with separate 2-way (dual-task type x set size) ANOVAs for each age group, revealed that this interaction was significant for older adults ($F[1, 16] = 7.01$, $\eta_p^2 = .31$, $p < .05$), but not for younger adults ($p > 88.$). A follow-up on the

2-way interaction for older adults with two separate 1-way (dual-task type) ANOVAs for set size 9 and 25 respectively, showed that there was a significant simple main effect of task type for the 9-item set size ($F[1,\ 16] = 15.28$, $\eta_p^2 = .49$, $p < .01$), but not for the 25-item set size ($p > 89.$). Dual-task complexity thus had no effect on performance of younger adults: Their performance was similar in both the condition with a relevant secondary task as well as the condition without a relevant secondary task. However, for older adults, MDEV performance did not change between the dual and the dual-r condition for the 25-item visual search display, but MDEV increased going from the dual to the dual-r condition when the 9-item visual-search display was presented. One explanation for these findings might be that older adults, with the higher set size, ignore the visual-search display, leading to similar results on the LCT. Evidence for this explanation comes from the high number of misses for older adults with an increase in set size.

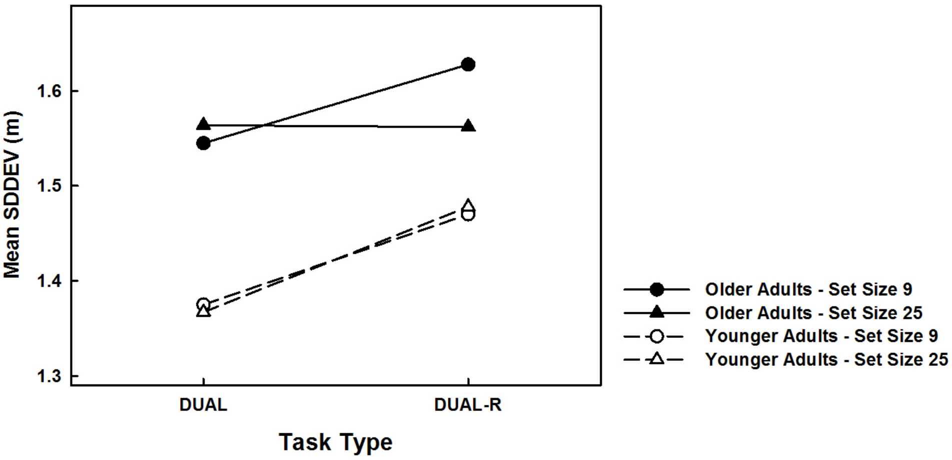


Figure 6.5: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25).

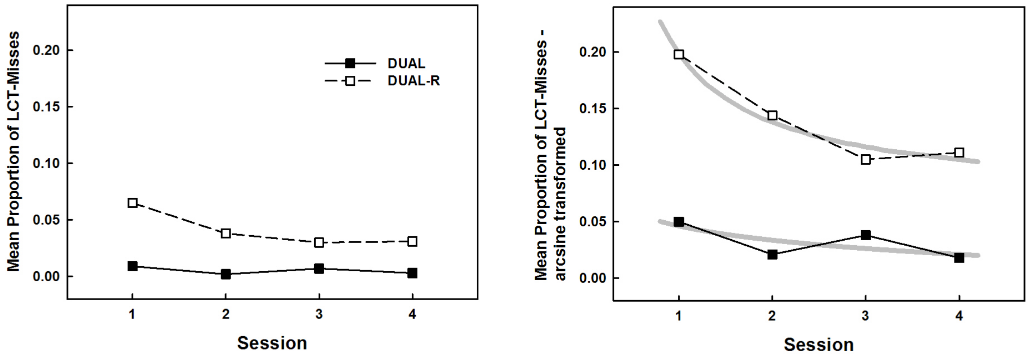
A similar pattern was found for SDDEV values. Figure 6.5 presents mean SDDEV as a function of dual-task type, age and set size. The factor dual-task type interacted significantly with the factor age ($F[1,\ 32] = 5.82$, $\eta_p^2 = .15$, $p < .05$), but this interaction was modulated by a significant 3-way interaction between the factors dual-task type, set size and age ($F[1,\ 32] = 5.51$, $\eta_p^2 = .15$, $p < .05$). To analyze this interaction, separate

2-way (dual-task type x set size) ANOVAs were conducted for younger and older adults respectively. The interaction was significant for older adults ($F[1,\ 16] = 7.77$, $\eta_p^2 = .33$, $p < .05$), but not for younger adults ($p > .60$). A further analysis of the significant interaction for older adults with separate 1-way (dual-task type) ANOVAs for each set size, revealed a simple main effect of dual-task type for set size 9 ($F[1,\ 16] = 16.26$, $\eta_p^2 = .50$, $p < .01$), but not for set size 25 ($p > .96$). As can be seen in Figure 6.5 the pattern for SDDEV values was thus the same as for MDEV values: Younger adults' SDDEV performance is not affected by dual-task type complexity. Older adults however, suffer from an increase in dual-task complexity as can be seen by an increase in SDDEV values, but only when the visual-search display with 9 items is presented. As for the MDEV values, these findings might be due to older adults ignoring the visual-search display when higher set sizes are presented.

## 6.3.4 Third Research Question

The third research question examines whether the use of a driving-task relevant secondary task has an effect on learning in dual-task driving situations.

One indication that the use of a driving-task relevant secondary task has an effect on learning in dual-task driving situations, comes from analysis on the proportion of LCT-Misses. Figure 6.6 presents the proportion of LCT-Misses as a function of session and dual-task type for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions. The ANOVA revealed a significant interaction between both factors ($F[3,\ 96] = 3.32$, $\eta_p^2 = .09$, $p < .05$). A post-hoc analysis with separate 1-way (session) ANOVAs for dual and dual-r, revealed a simple main effect of session for dual ($F[3,\ 99] = 3.21$, $\eta_p^2 = .09$, $p < .05$) as well as dual-r conditions ($F[3,\ 99] = 5.93$, $\eta_p^2 = .15$, $p < .01$). Practice thus had a positive effect on the proportion of LCT-Misses in both the dual as well as the dual-r condition. An illustration of this effect by means of learning curves shows that the increase in performance was strongest for the dual-r condition (fitted power function: $y = 0.08 - 0.12x^{-1.02}$, $R^2 = .96$, RMSE $= 0.00$) as compared to the dual condition (fitted power function: $y = 0.03 - 0.00x^{-0.00}$, $R^2 = .53$, RMSE $= 0.13$). As can be seen in this example and especially for the dual-condition, power laws are not always adapted to describe learning progress.



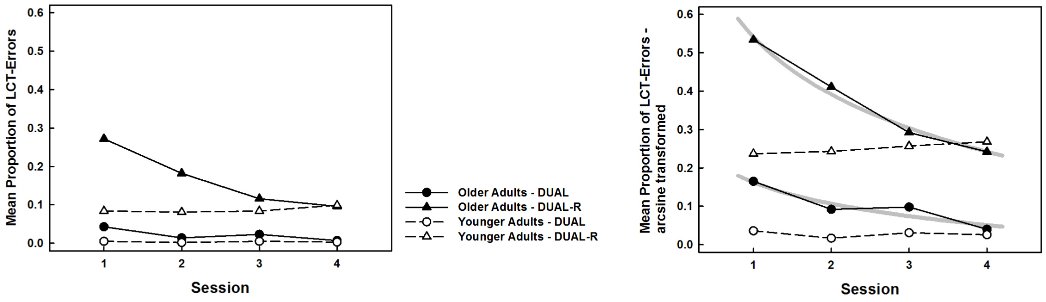
(a) (b)

Figure 6.6: Proportion of LCT-Misses as a function of session (1-4) and dual-task type (dual, dual-r) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions.

### 6.3.5 Fourth Research Question

The final research question examines whether the use of a driving-task relevant secondary task has an effect on learning between age groups.

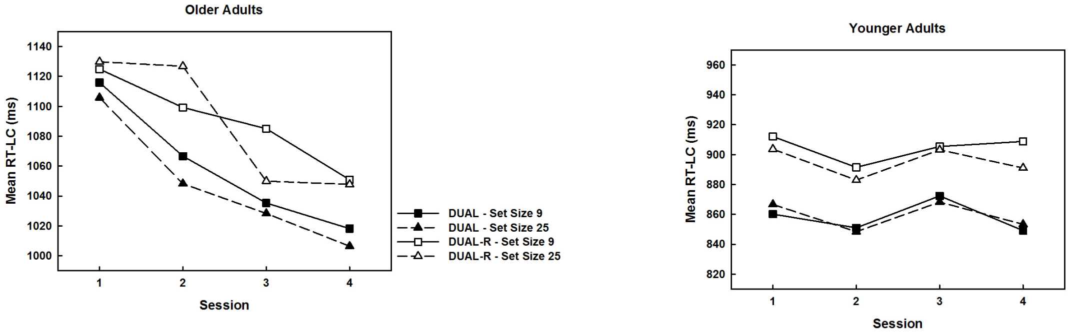
A first indication that the use of a driving-task relevant secondary task has an effect on learning between age groups, comes from LCT-Errors. Figure 6.7 presents the mean proportion of LCT-Errors as a function of session, age, and dual-task type for (a) non-arcsine transformed values and (b) arcsine-transformed values, with fitted power functions for the older age group. The factor session interacted significantly with the factor age ($F[3,\ 96] = 15.54$, $\eta_p^2 = .33$, $p < .001$), but this 2-way interaction was modulated by a significant 3-way interaction including the factors session, age and dual-task type ($F[3,\ 96] = 3.51$, $\eta_p^2 = .10$, $p < .05$), represented in Figure 6.7. A follow-up analysis of this interaction with separate 2-way (session x age) ANOVAs for the dual and dual-r condition respectively, revealed a significant interaction between session and age for the dual condition ($F[3,\ 96] = 3.85$, $\eta_p^2 = .11$, $p < .05$) as well as for the dual-r condition



(a) (b)

Figure 6.7: Proportion of LCT-Errors (a) session (1-4), age (younger, older) and dual-task type (dual, dual-r) for (a) non-arcsine transformed and (b) arcsine-transformed values, with fitted power functions for the older age group.

($F$[3, 96] = 8.59, $\eta_p^2$ = .21, $p$ < .01). Follow-ups of each 2-way (session x age) ANOVA with separate 1-way (session) ANOVAs for younger and older adults respectively and for each dual-task type condition, revealed for the dual condition a simple main effect of session for older adults ($F$[3, 48] = 4.60, $\eta_p^2$ = .22, $p$ < .05), but not for younger adults ($p$ > .46). In the dual-r condition, there was a simple main effect of session as well for older adults ($F$[3, 48] = 9.73, $\eta_p^2$ = .38, $p$ < .01), but not for younger adults ($p$ > .66). Practice thus does not have an effect on the proportion of LCT-Errors for younger adults in both the dual and the dual-r condition. For older adults however, practice has a positive effect and the proportion of LCT-Errors decreases over time for both the dual as well as the dual-r condition. As illustrated by the learning curves, the decrease in LCT-Errors is stronger for the dual-r (fitted power function: $y$ = 0.54 - 0.14$x^{-0.73}$, R$^2$ = .99, RMSE = 0.03) condition as compared to the dual condition (fitted power function: $y$ = 0.16 - 0.06$x^{-0.58}$, R$^2$ = .90, RMSE = 0.02).



(a) (b)

Figure 6.8: Mean Reaction Time until Lane Change (RT-LC) as a function of session (1-4), dual-task type (dual, dual-r) and set size (9, 25) for (a) older adults and (b) younger adults.

Another indication that dual-task complexity (i.e., dual versus dual-r) influences learning effects between age groups comes from data on lane-change onset times (RT-LC). Figure 6.8 presents mean RT-LC as a function of session, dual-task type and set size for (a) older adults and (b) younger adults. The factor set size showed a tendency but missed to reach significance ($p$ = .09). The 4-way interaction between those factors however, reached significance ($F[3, 96] = 3.26$, $\eta_p^2 = .09$, $p < .05$). To analyze this interaction, two separate 3-way ANOVAs including the factors session, dual-task type and set size were conducted for each age group. For the younger adults the 3-way interaction session x dual-task type x set size was not significant ($p > .40$). For the older adults, this same interaction was significant ($F[3, 48] = 3.43$, $\eta_p^2 = .18$, $p < .05$). A further post-hoc analysis of this significant 3-way interaction with separate 2-way (session x setsize) ANOVAs for the dual and the dual-r condition respectively, revealed no significant interaction for the dual condition ($p > .62$) nor for the dual-r condition ($p > .21$). The interaction shown in Figure 6.8a can thus be explained as follows: RT-LC values for younger adults do not change over sessions. RT-LC values for older adults however decrease over sessions. Although the 2-way interactions (session x set size) were not significant, patterns differ for the two task types (dual, dual-r), as revealed by the significant 3-way interaction including the factors session, dual-task type and set size. As can be seen in Figure 6.8a, RT-LC values decrease more or less in parallel for both set sizes in the dual condition, whereas in the dual-r condition, RT-LC decreases steadily for the 9-item set size, but not for the 25-item set size. These findings indicate that dual-task complexity has no effect on learning for younger adults. For older adults however, some differences in learning can be observed according to the dual-task condition. As such, learning rates in the dual condition (i.e., without relevant secondary task) do not differ a lot, even with different set sizes. However, in the dual-r condition, secondary-task complexity seems to play a role leading to differences in learning rates for different set sizes.

## 6.4 Discussion

The first research question examined whether secondary-task relevance influences performance in dual-task driving situations. The results of this experiment show the following findings: The number of LCT-Errors (i.e., when the participant either changed into the wrong direction or the wrong number of lanes) as well as the number of LCT-Misses (i.e., when the participant was supposed to stay on his lane, that is drive straight on, but changed lane instead) increased with the relevant secondary task as compared to the condition with the irrelevant secondary task. Those findings were backed up by data concerning lane-change accuracy (as witnessed by MDEV and SDDEV values): Lane-change accuracy decreased when the secondary task was of relevance for the driving task. These findings are partially in accordance with findings by Seppelt and Wickens (2003), who found, based on the multiple-resource theory (Wickens, 1984), that a relevant visual messages interfered more with the driving task than an irrelevant visual message. However, those same authors also found that relevant messages had less impact on the driving performance than irrelevant messages. One explanation for the difference in findings between this study and the study of Seppelt and Wickens (2003) is the nature of the secondary task. In Seppelt and Wickens' experiment, the secondary task was a visual or vocal message, which participants only needed to repeat. No sensory-motor response (i.e., button press) on the message was expected and the message could even be ignored without affecting the driving task. In the setting of this experiment, the secondary task was a visual search task, where not only participants needed to perform a cognitive effort to solve a problem, but which furthermore defined the upcoming driving instructions. In other words: There was a causal relationship between the answer on the secondary task and the driving task instructions for the upcoming lane-change. As opposed to Seppelt's experimental setting (Seppelt & Wickens, 2003), the secondary task in this experiment could not be ignored without affecting the driving task. The observed drops in driving performance measures could therefore be due to the fact that participants were forced to divide attention between the driving and the secondary task, leading eventually to declines in driving performance as compared to a dual-task situation in which the secondary task could be ignored.

There was no main effect of relevance on secondary task measures. A more detailed analysis showed though that these performance measures also depended on the age of the participant and the secondary task difficulty. Details of these findings will be discussed in the paragraphs that follow. This experiment showed however, that a relevant secondary task caused more interference (as witnessed by IRI values) than an irrelevant secondary task. This effect was especially strong when complexity of the secondary task increased (i.e., with higher set sizes). The use of a relevant secondary task, in the form of a visual search task which was of relevance for the next lane-change maneuver, thus had an effect on driving performance measures: Performance was worse with a relevant secondary task as compared to the condition with an irrelevant secondary task.

As to the second research question, whether performance differed between younger and older adults with a secondary task of relevance for the driving task, the findings of this experiment showed that performance differences between age groups exist. A first indication comes from interference values (as expressed by the IRI variable): Like in the previous experiment, interference from the secondary task was higher for older adults than younger adults. The difference between groups was however stronger when a relevant secondary task was used as compared to an irrelevant secondary task. This finding is as expected: Older adults need more time discerning a target amidst distractor items (Hahn et al., 2011) leading to longer search times on the VST. When the VST is of relevance for the driving task (i.e., by providing information for the next lane-change maneuver) differences between older and younger adults might become more obvious in reaction time measures, as the VST cannot be ignored (as is the case when the secondary task is irrelevant for the driving task). Longer reaction times on the VST, leads to stronger interference in general, hence resulting in higher IRI values. These explanations are backup up by literature on processing styles for both older and younger adults. Previous research has shown, that older adults generally engage in more serial processing of tasks (Göthe et al., 2007). Indeed in Experiment 3, older adults tended to first perform the driving task and even ignore the secondary task if needed (leading to a high proportion of misses). The driving task was correctly prioritized over the secondary task by providing feedback and explicit task instructions. However, in the experimental setting of this experiment, not only the driving task was supposed to be prioritized (and reinforced with feedback and task instructions), but the secondary task was of importance as well, as it provided important information for the next lane-change maneuver. In a certain way, older adults were therefore "forced" to process in a parallel manner, in order to on one hand respect the task constraint that the driving task was the primary task, but on the other hand perform the secondary task to obtain information for the next lane-change maneuver. As seem to indicate the data, older adults, experiencing difficulties in parallel processing (Göthe et al., 2007), seem to have chosen to shift their task priority towards the secondary task, which might eventually have led to the observed drops in driving performance measures.

A further indication that the driving performance of older adults suffered from task relevance, comes from lane-change accuracy data. Relevance of the secondary task had no effect on lane-change performance of younger adults, whereas lane-change performance of older adults decreased significantly when the secondary task was relevant as compared to the condition in which it was irrelevant. This effect was influenced by secondary task complexity though (i.e., the number of items on the visual search display): With the 25-items set size, no difference in performance could be found in both MDEV as well as SDDEV values between the condition with a relevant secondary task as compared to the condition with an irrelevant secondary task. However, with the 9-items set size, lane-change accuracy of older adults deteriorated in the condition with the relevant secondary task. This might actually be an indication that priority shifts towards the relevant secondary task, affected driving performance. An explanation for the fact that those findings were not replicated with the 25-items set size, might come from the number of misses on the VST: With the larger set size, the number of misses on the VST increased significantly as compared to the condition with the smaller set size.

Older adults might have been overwhelmed with the relevant secondary task including 25 items, which eventually might have led them to ignore the VST. This, in turn, might however have had a positive effect on driving performance measures.

The third research question examined whether the use of a driving-task relevant secondary task had an effect on learning in dual-task driving situations. The current results replicate findings of the previous experiment and confirms that practice had a positive effect on learning: The proportion of LCT-Misses (i.e., when the participant was supposed to stay on his lane, but changed lane instead) decreased over sessions. This finding is in accordance with the ACT-R theory by Anderson (1982). However, this decrease was strongest when the secondary task was relevant for the driving task. Although this finding is not in accordance with initial expectations (i.e., learning was expected to be harder in the relevant condition, leading to less steep learning curves as compared to the irrelevant condition) looking at Figure 6.6, the current finding can be explained as follows: As performance in the relevant condition is worse at the beginning, a stronger learning effect can be observed as compared to the irrelevant condition. In the irrelevant condition, participants quickly reach a ceiling effect, leaving little to no room for strong improvements. This results in a more moderate learning curve as compared to the relevant condition.

The final research question examined whether the use of a driving-task relevant secondary task had an effect on learning between age groups. This experiment provides some indications that the use of a driving-relevant secondary task has an effect on learning between age groups. A first indication comes from the proportion of LCT-Errors produced (i.e., when either the participant changed into the wrong direction or the wrong number of lanes): Performance on this measure did not differ for younger adults in either the relevant or the irrelevant condition. However for older adults, practice had a positive effect and the proportion of LCT-Errors decreased with practice. This decrease was strongest in the relevant condition. Although no learning took place for the younger adults, the observed differences in learning between younger and older adults were as expected: Due to age-related differences in cognitive processes needed for learning (Anderson, 1982), for solving visual search tasks (e.g., working memory functions needed to discern relevant from non-relevant information; Hahn et al., 2011) and managing dual-task situations (Wilschut et al., 2008), older adults were expected to learn at a slower rate than younger adults. Again, the difference in learning rate between the relevant and the irrelevant condition can be explained by the fact that the proportion of LCT-Errors was much higher in the relevant condition as compared to the irrelevant conditon, therewith leaving more room for improvement. As to the strong differences between the relevant and the irrelevant condition, the nature of the secondary task might have played a role. In the relevant condition, the secondary task needed to be solved to obtain instructions for the oncoming lane-change maneuver. In other words: Participants were encouraged to process in a parallel manner. Especially older adults have difficulties with parallel processing of information (Göthe et al., 2007), leading to potential drops in performance.

Another indication provided by this experiment that relevance of the secondary task influences learning effects between age groups, comes from lane-change onset times. Again, no improvement can be observed for the younger adults with practice. This finding can be explained by the younger adults being on a high level of performance from the first session on, leaving little to no room for improvement. However, for older adults, lane-change onset times decrease with practice. Secondary-task relevance leads to differences in learning curves: Learning rates are strongest in the condition with a relevant secondary task. This effect is even stronger with bigger set sizes. These latter findings are as expected, as age effects impair discerning a target among distractor items (Hahn et al., 2011). Higher set sizes provide more distractors, making it harder to detect the target. The observed difference in learning rate between the relevant and the irrelevant condition can again be explained by the nature of the secondary task, as explained in the previous paragraphs.

## 6.5 Conclusion

With this experiment, the effect of a relevant secondary task on learning differences between younger and older adults in a dual-task driving situation was examined. This experiment differed from previous experiments, in that a visual search task was used that needed to be attended to in order to obtain information for the driving task. It was expected that this type of secondary task would lead to a lower performance on all measures, especially for older adults. Based on the findings of Experiment 3 though, learning effects were expected to have a positive effect on dual-task performance for both younger and older adults. Taking into account aging effects, it was expected that older adults would benefit a little bit less from practice than younger adults though.

Taken together, the results of this experiment show that a relevant secondary task which cannot be ignored had a strong effect on driving performance measures: Driving efficiency dropped under the influence of a relevant secondary task. This experiment showed as well, that especially older adults suffer from this type of relevant secondary task, leading to performance losses in a dual-task driving situation. Learning had a positive effect on performance of both age groups though: Especially driving performance measures improved with practice over sessions. This effect was especially strong for older adults and with the relevant secondary task. These findings show that with the proper amount of training, older adults are capable of improving their performance in difficult dual-task situations where they are encouraged to divide their attention between two tasks that are (almost) equally demanding.

# 7 General Discussion

This final chapter is divided into four sections. First, the major findings of this thesis as well as their accompanying interpretations are summarized. Then, the general limitations concerning all experiments are mentioned and commented. Third, practical relevance of this research for the use of IDSS in vehicles are addressed and finally a number of future directions.

## 7.1 Summary of Findings and Interpretations

Previous research has shown that driving is a complex task, combining the analysis of continuous information input (Jackson et al., 2013) with coordinated motor output to assure a safe maneuvering of the vehicle (Anstey et al., 2005). Driving can indeed be considered multi-tasking, constituted of different subtasks, operating at different levels of priority and complexity (Michon, 1985). Of particular importance for safety are decisions/actions at the operational level, as they are immediate reactions to (potentially dangerous) traffic situations, operating in the millisecond to second range (e.g., emergency braking). Research has shown that especially older drivers have sometimes difficulties handling these types of situations which require quick planning, problem solving and decision making (Anstey et al., 2005). These difficulties are due to age-related changes in for example working memory (Borella et al., 2008) and executive functions (Adrian et al., 2011) needed for discerning important from unimportant information (Hahn et al., 2011), planning (Allain et al., 2005; Sorel & Pennequin, 2008), problem solving (Diehl et al., 1995) and decision making (Henninger et al., 2010). Driving assistance at this level can be helpful, as it will help the driver to make decisions or prepare for better decision making (or even take over completely), which will reduce reaction times, leading to potentially safer driving. Research has however shown that often IDSS initially lead to negative behavioral adaptation. Examples on studies with Adaptive

Cruise Control (ACC) have shown that initial use can for example lead to higher speed, smaller minimum time headways and larger brake forces (Hoedemaeker & Brookhuis, 1998, see the *Risk Homeostasis Theory* as well; Wilde, 1989). Other authors observed general reductions in the driver's level of attention and shifts in attention away from the road scene (Brookhuis et al., 2001; Gruendl, 2005; Ranney, 2008)

Wilschut (Wilschut et al., 2008; Wilschut, 2009) showed that especially older adults suffer from dual-task situations, leading to drops in tracking performance on a driving-like task and increases in reaction times on a visual search task. Indeed, cognitive processes needed for dual-task performance (e.g., memory and attention processes; Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) decline at later age. Taking into account models of learning however, it can be assumed that those initial difficulties can be overcome with practice. As such, according to the Adaptive Control of Thought − Rational (ACT-R) theory of skill acquisition by Anderson (Anderson, 1982), practice will lead to triggering automatic problem-solving routines, demanding few cognitive resources and allowing to do other activities in parallel. A similar proposition comes from the Skill-Rule-Knowledge Model (SKR-Model) by Rasmussen (1983), according to which human beings go from knowledge-based behavior to skill-based behavior with practice. At this level, sensory-motor performances become automated and require very little or no conscious control, freeing up cognitive resources (Wickens, 1984). Based on these models, it can be expected that initial difficulties experienced when driving with IDSS can be overcome with practice. Learning however, heavily depends on memory and attention processes (Repovs & Baddeley, 2006), which, as mentioned above, decline at older ages (Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010).

The purpose of this thesis was therefore to explore a particular question within the research domain of dual-task driving: *Can practice help older people to drive with IDSS more successfully?* To our knowledge, hardly any study looked at the effect of aging on learning in specidif dual-task driving situations, that is driving with IDSS. It thus remains to be explored whether practice could help improve performance of both younger and older adults in a dual-task driving situation. Based on previous research on learning for both older and younger adults, it was expected that practice would lead to performance improvements in a dual-task driving situation. However, due to age-related differences especially in cognitive processes, it was expected that older adults would experience more difficulties in learning than younger adults.

While previous attempts by other authors at (partially) answering similar research questions, were difficult to interpret due to methodological issues (e.g., Chisholm et al., 2008; Cooper & Strayer, 2008; Shinar et al., 2005), a controlled laboratory setting was used, with a simulated, simplified driving environment (LCT; Mattes, 2003) as the primary driving task and a visual search task (Treisman & Gelade, 1980) as a surrogate IDSS (for more information on the exact experimental procedure, see the General Method, Chapter 2).

### 7.1.1 Experiment 1: Methodological Study

In Experiment 1, it was examined whether the analysis of lane-change performance on the LCT could be adapted to individual driving styles allowing a more precise analysis of individual driving performance. The aim was exploring (a) whether the use of a relative calculation method could be used for defining relevant segments (more difficult versus easier) within the LCT, (b) whether segments based on this relative calculation method were better adapted to individual lane-change behavior and finally (c) whether the use of a relative calculation method for the definition of both segments would more precisely reflect age differences in lane-change performance. Based on individual lane-change behavior, the driving task was divided into lane-change and lane-keeping segments. The effect of those segments was measured by analyzing performance data of younger and older adults on the LCT.

Apart from examining whether the use of a relative calculation method could be used to define lane-change and lane-keeping segments, performance metrics with relative LC and LK windows were expected to be better adapted to actual lane-change behavior and therefore yield more accurate performance measures. This approach was furthermore expected to be more sensitive to age-related changes, like for example in the sensory-motor system (Alm & Nilsson, 1995; Jagacinski et al., 1995; Rubichi et al., 1999; Shanmugaratnam et al., 2010).

Findings in Experiment 1 showed indeed that the kinematics of the vehicle's displacement can be used to define the start as well as the length (and therewith the end) of each lane-change maneuver. This method allows defining relative calculation windows (LC and LK) taking into account individual driving style of each participant. In contrast to absolute windows (in which the segments have a fixed length for all participants), relative windows are defined based on each driver's individual lane-change behavior, leading to more representative (i.e., better adapted to the individual driving style of each participant) segmentations. Relative windows furthermore turned out to be more precise and more representative of individual lane-change behavior, hence better reflecting age differences as well. Indeed, due to a general slowing in tracking tasks at older age (Jagacinski et al., 1995), as well as longer reaction times (Alm & Nilsson, 1995; Rubichi et al., 1999; Shanmugaratnam et al., 2010), older adults' lane-change maneuver were generally longer than those of younger adults. With the relative calculation, this difference was taken into account, leading to segments that were well adapted to the individual lane-change length, better reflecting differences in performance between age groups than when using absolute segments. By calculating performance values based on actual driving behavior, the effect thus got less attenuated by the inclusion of erroneous data points. In contrast: When using absolute segments, driving performance values were potentially over- or underestimated due to incorrect segmentation borders. Another advantage of the relative calculation method is the creation of two additional variables to the "standard" LCT variables: Reaction Time until Lane Change (RT-LC) and Movement Time (MT-LC). Both variables provide valuable insight into lane-change behavior. RT-LC examines how long it takes a participant to start the actual lane-change maneuver. RT-LC has the advantage to serve as an event-detection measure providing a reaction time to a precise event (Harbluk et al., 2007; Young et al., 2011). MT-LC represents the period of time in which a participant is actually changing lanes. It has the advantage of being independent of the optimal track defined by the LCT.

### 7.1.2 Experiment 2: Effect of Feedback on Dual-Task Performance

A second methodological issue that needed to be clarified was how to make sure participants would prioritize the driving task, despite being in a simulated driving environment (which, unlike real-life driving situations, can lead to priority shifts towards the secondary task; Levy & Pashler, 2008). End-of-block feedback was provided (Summary Knowledge of Results (SKR); Schmidt et al., 1990) in addition to explicit instructions to prioritize the driving task over the secondary task. The second experiment was aimed at exploring whether driving performance feedback on the LCT in the form of SKR actually helped participants to prioritize the driving task over the secondary task. It was furthermore examined whether younger and older adults equally benefit from feedback on the LCT. Finally, as feedback is an important tool for learning as well, as it provides guidance (Salmoni et al., 1984; Schmidt et al., 1990) and motivation (Vollmeyer & Rheinberg, 2000), it was also investigated whether feedback on the LCT had an effect on learning, potentially resulting in improved performance on the LCT, the secondary task and in the dual-task condition. A group of younger and older adults had to perform the LCT and a visual secondary task. Half of the participants received feedback about their LCT performance in the form of SKR and the other half did not receive any feedback about their LCT performance.

Based on previous research (Brumby et al., 2007), it was expected that drivers with feedback would apply prioritization strategies to allow for a better control of attention between both tasks in favor of the driving task, as compared to drivers who did not receive any feedback regarding their driving performance. However, as cognitive resources responsible for the distribution of attention between two or more tasks (Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) as well as their prioritization are affected by age-related changes (Borella et al., 2008; Sweeney et al., 2001), differences in the allocation of attention between younger and older adults were expected. More precisely, it was expected that older adults, under the effect of feedback, would allocate less attentional resources to the secondary task and concentrate more on the primary task. For this reason, a greater loss of performance on the secondary task for older adults with feedback was expected, as compared to younger adults with feedback. As to investigated learning effects, the group of drivers with feedback were expected to show larger improvements in their driving performance with practice over blocks due to enhanced motivation (Vollmeyer & Rheinberg, 2000) and a more consistent improvement of performance (i.e., variability in performance will be reduced; Berthon-Donk et al., 2011; Schmidt et al., 1990), as compared to the group without feedback.

Results in Experiment 2 provided evidence that SKR feedback was a useful tool to prioritize the driving task in a dual-task paradigm including the LCT and a visual search task. Indeed, several driving performance measures improved under the effect of feedback, providing an indication that in a dual-task situation, under the effect of feedback, attention was turned in priority to the driving task. Another indication that indeed feedback had a positive effect on task prioritization comes from the number of misses produced on the secondary task by older adults, which was higher when feedback was provided. This indicates that when interference from a secondary task occurs, staying focused on the primary task results in performance reductions on the secondary task. These latter findings are in accordance with the Attentional Resource Theory of Kahneman (1973) according to which a drop of performance will be observed when resources for one or both tasks are reduced due to capacity limits of the system. As expected, the observed effect of feedback on secondary task performance was stronger for older adults, as due to aging effects, the attentional resources responsible for dividing attention between tasks are less effective (Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010). These findings are not only in accordance with findings from previous studies (Brumby et al., 2007; Horrey et al., 2006), but prove furthermore, that even with feedback on both the LCT performance as well as the secondary task performance (i.e., instead of on one performance variable only; Brumby et al., 2007), priorities are respected if participants are explicitly instructed to do so. A bit unexpected, no benefit of feedback on learning effects was found. On the contrary: The group without feedback produced more accurate driving performance outcomes. The current data suggest, that these findings might be due to differences in motivation (Vollmeyer & Rheinberg, 2000): The group with feedback was apparently so motivated that they would more or less immediately reach their maximum performance, leaving little to no room for improvement.

### 7.1.3 Experiment 3: The Effect of Aging on Learning in Dual-Task Driving Situations

Experiments 1 and 2 provided a new methodological basis for conducting two learning studies. The goal of the first learning experiment was to investigate the effect of aging on learning in dual-task driving situations. The motivation of this study was the observation that on one hand, older adults suffer more from an additional task to driving (e.g., Wilschut et al., 2008), but on the other hand indications exist that learning might help overcome those initial difficulties. Based on the observation that more cognitive resources (Wickens, 1984) become available with practice due to the (partial) automatization of subtasks (Anderson, 1982), a positive effect on performance can be expected. Especially taking into account Kahneman's attentional resource theory (1973), it can be expected that dual-task interference will become less with practice, as more resources become available. To account for effects of practice and learning, Experiment 3 consisted of 4 sessions of approximately 3 hours each. The first three sessions were aimed at studying the effect of repeated practice on performance, the fourth session served as a retention session, which took place at a later date (for more information on the experimental design, see the General Method Section (Chapter 2.3).

Practice was expected to have an effect on both driving as well as secondary task measures. According to the Skill-Rule-Knowledge-Model of Rasmussen (SKR-Model 1983), with increasing experience, information will be increasingly processed on a skill-based level, where sensory-motor performance becomes automated and requires very little or no conscious control. It was therefore expected that driving performance measures would improve with practice, therewith freeing up resources, allowing more attention to be directed to the secondary task, hence improving this task as well. To specifically measure dual-task interference, a new variable was defined, Inter-Response Interval (IRI; for a more detailed description, see the Dual-Task section of the General Method, Chapter 2.4.3). As practice was expected to have a positive effect on automation of the driving task, there should be more resources available to perform the secondary task, leading to less dual-task interference (i.e., shorter IRI lengths). As a complement to our research question, retention effects were studied to examine whether discontinuation of practice had an effect on performance. In line with findings by Kantak and Winstein (2012), decay effects were expected, as according to an extensive literature review by those authors, the largest part of skills is forgotten one week after the last training session.

Skill acquisition, as well as retention of this skill, heavily depend on memory processes (Wagner, 2006), which are known to decline at later ages (Borella et al., 2008). Therefore, differences in learning and retention between younger and older adults were expected. Indeed, previous studies had shown (for example Caird et al., 2008; Göthe et al., 2007; Verwey et al., 2011) that age-related differences are likely to persist, even after extended practice. Older adults were therefore expected to never reach the same level of performance as younger adults, despite a benefit of practice. Due to an age-related decline in cognitive processes (Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) that underlie skill acquisition (e.g., attentional processes, memory functions; Anderson, 1982) and which result in a general slowing of this process, older adults were furthermore expected to learn at a slower rate than younger adults.

Results of Experiment 3 showed that practice had a positive effect on dual-task driving performance for both younger and older adults. Performance in both the driving as well the secondary task improved. Dual-task interference decreased as well, as witnessed by smaller IRI values. These findings are in accordance with Anderson's ACT-R theory (1982) as well as Rasmussen's SKR-Model (Rasmussen, 1983), according to which tasks become (partially) automated with practice, thereby freeing up resources, leading to improvements on all performance measures. As to the effect of aging on learning, results were partially as expected: Although older adults benefited from practice, their maximum performance reached the same level as untrained younger adults. This is in accordance with findings by other authors (for example Caird et al., 2008; Göthe et al., 2007; Verwey et al., 2011). Differences in learning between age groups can be explained by age-related slowing in cognitive processes (Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) such as attentional processes and memory functions, which are important for skill acquisition (Anderson, 1982). In fact, our data even show that older adults benefit less from practice in more difficult driving situations (i.e., lane-change segments). These findings are unexpected, but one explanation could be that the dual-task situation was so complex for them, that even with practice they could not free enough resources to improve their dual-task performance. An important finding of our experiment, was that the retention effect had no effect on performance. Acquired skills in a dual-task driving environment thus seem to remain stable over time, despite lack of practice. It might furthermore indicate that the three sessions used in our experiment were sufficient to establish a stable learning outcome (insensitive to decay) for both younger and older adults.

### 7.1.4 Experiment 4: The Effect of Aging on Learning in Relevant Dual-Task Driving Situations

In extension to Experiment 3, in Experiment 4 a dual-task study was conducted aimed at examining age effects on practice in a dual-task driving environment, but in which the secondary task was of relevance for the driving task. The main motivation for this setting was that often in dual-task driving studies, the primary and the secondary task are not related. However, in a real-life driving context, ignorance of the IDSS by the driver might lead to discomfort (e.g., when navigation instructions are ignored) or even to dangerous situations (e.g., when warnings of an Advanced Cruise Control are ignored). Although some research in that direction had been done (see Seppelt & Wickens, 2003), with our experiment both younger and older adults were pushed to their limits by providing a secondary task which could not be ignored, as it provided direct instructions for the driving task. In other words: Ignorance of the secondary task would lead to performance loss on the driving task. Particularly, the following research questions were examined: (a) Does a relevant secondary task influence performance in dual-task driving situations? (b) Does performance differ between younger and older adults with a relevant secondary task? (c) Does the use of a relevant secondary task have an effect on learning in dual-task driving situations? (d) Does the use of a driving-task relevant secondary task have an effect on learning between age groups?

To account for effects of aging on practice and learning in a situation where the secondary task is of relevance for the primary task, the experiment consisted of 4 consecutive experimental sessions, in which in addition to the standard tasks (for more information on the experimental setting, see Chapter 2.3), participants also had to perform a dual-task condition, in which the secondary task was of relevance for the driving task (dual-r condition, for detailed information see Chapter 6.2.2).

According to the Multiple-Resource-Theory (Wickens, 1984), it was expected that a relevant visual search task would interfere more with the visual driving task than an irrelevant visual search task. Visual attention was expected to be directed off the road to attend to the visual search task, which was expected to have a negative effect on driving performance measures. These negative effects were expected to be stronger for older adults as compared to younger adults: As a visual search task demands discerning a target item in between distractors, for which older adults need more time (Hahn et al., 2011), differences in performance were expected to become more obvious between age groups. According to the ACT-R theory by Anderson (1982), improvements in the dual-r condition with practice were expected. However, as the relevant secondary task was more distracting than the non-relevant secondary task, learning curves were expected to be less steep in the condition in which the secondary task was of relevance.

The results in Experiment 4 indeed showed that a relevant secondary task, which cannot be ignored, had a strong effect especially on driving performance measures, particularly for older adults. As expected and in accordance with Wickens' theory (1984), attention was forced off the road to the visual search task, leading to a negative effect on driving performance measures. Learning had a beneficial effect on both age groups though: Driving performance improved with practice over sessions. Especially older adults benefited from practice, by increasing their capacity to divide their attention between the primary and the secondary task, which were now almost equally demanding. This led to higher performance on both the primary as well as the secondary task.

## 7.2 Limitations of This Study

Although uttermost care has been taken throughout this thesis to respect scientific methods and practices, some limitations must be mentioned. A first limitation is the laboratory environment using the LCT as a driving task and a visual search task as a surrogate IDSS. The degree to which findings can be generalized is limited by the trade-off between ecological validity and experimental control of any experiment (Loomis, Blascovich, & Beall, 1999). According to Hofmann (2011, p. 151), "task design is a compromise between the needs of experimental control and direct applicability to real driving". From an applied perspective though, some factors reduce the ecological validity of this study.

As mentioned before, a first aspect which might have reduced the external validity of our experiment is the use of the LCT as the primary driving task. The LCT is a highly simplified driving environment, consisting of a straight three-lane road without any other traffic and in which participants need to maintain a constant speed to control for compensatory strategies, like for example reducing speed when secondary task demand increases (for a more detailed description of the LCT, see Chapter 2.2.1 in the General Method). Through its simplified nature, the LCT allows for a highly controlled driving task, which is simple, reliable (Benedetto et al., 2011) and sensitive to secondary task demand (Bruyas et al., 2008; Harbluk et al., 2009; Maciej & Vollrath, 2009), but its data should also be interpreted with caution. The LCT, due to its nature, is more of a tracking task in which participants have to react as fast as possible to a traffic sign. The reference track (i.e., the "basic" reference track) is rather artificial. An argument in favor of using the LCT despite this objection, is that when the same reference track for all experimental conditions is used, it can still be useful for between-group comparisons (see for example Berthon-Donk et al., 2011) or between-condition comparisons (see for example Harbluk et al., 2007). In the current experiments, in which between and/or with group conditions were compared, the LCT has proven to be a useful tool, providing important first insights into learning behavior from both younger and older adults with different task complexities. Nevertheless, the rather artificial LCT cannot be compared to studies using field tests or more sophisticated driving simulators. Further research in more naturalistic environments is needed to validate our findings (although research has shown that results obtained in a driving simulator often hold for real-life driving as well; Carsten & Brookhuis, 2005).

Another point of criticism to the LCT is that the basic reference track does not take into consideration driving behavior of different age groups: Reaction times until lane-change start are much longer for older adults as compared to younger adults and they then take more time to actually change of lane, leading to even bigger deviation from the optimal reference track. Because of such limitations of the LCT standard analysis, additional analyses were conducted using relative calculation methods for segment definition: Every segment was adapted to each participant's individual lane-change style. Reaction-time until lane change was taken into account as well and served as a separate variable (see Experiment 1, Chapter 3).

Finally, previous studies have criticized that task instructions play an important role on LCT performance measures (Young et al., 2011): In the LCT ISO standard (ISO 26022, 2010) it is suggested to instruct participants to change lanes as soon and as quickly as possible after the lane-change information appears on the sign. However, no instructions are given concerning the completion of the lane change (i.e., before the next lane-change sign). Young and colleagues (2011) showed that a lack of those instructions led to more gradual and naturalistic lane changes, but potentially affected the level of attentional demand required by the LCT, making it less sensitive to differential effects of the secondary task. In our experiments, feedback was used to alter prioritization schemes, resulting in successfully prioritizing the primary task over the secondary task (see Experiment 2, Chapter 4). Nevertheless, a difference in instructions with other studies (especially with regard to lane-change length), might have led to potential performance differences.

In the same line of criticism, the visual search task used can be criticized as well as being too artificial (i.e., as opposed to real-world situations in which the target item is among other items arranged irregularly on what is usually a non-homogeneous background; Wolfe, 1994) and not representative of real-life IDSS (i.e., IDSS displays are rarely static, but update information continuously; Wilschut, 2009). Nevertheless, a visual search task has the advantage of providing a highly controlled task which is comparable with some (but certainly not all) IDSS. A navigation system for example demands a cognitive filtering of relevant from irrelevant information. Further research with real or close-to-real IDSS are necessary to validate findings of this study though.

## 7.3 Practical Relevance

This research resulted in improved insights into learning capabilities of younger and older adults in dual-task driving situations with different types of secondary tasks (i.e., relevant and irrelevant for the driving task) and clarified some methodological issues which could help future research assure even better methodological control when using the LCT.

On a methodological level, it is recommended to implement *feedback on the primary driving task when using the LCT* with a secondary task for which participants receive feedback as well. This allows participants to stay focused on the primary driving task, leading to more ecologically valid prioritization schemes (i.e., like in real-life driving). Furthermore, it is recommended to analyze the LCT data separately for more difficult and easier segments (i.e., lane-change versus lane-keeping segments), *based on individual driving behavior of the participant.* As shown by Experiment 1, especially when comparing age groups, this leads to a more representative segmentation of the driving task, which in turns leads to more accurate driving performance measures.

The experiments in this thesis showed that learning indeed had a positive effect on managing dual-task driving situations: Performance on both the primary as well as the secondary task improves and dual-task costs decrease with practice. Therefore, a *practice period* can be very helpful to overcome initial difficulties in such situations (e.g., when driving with IDSS) leading to better performance on the driving task. Taking into account the Three-Level Task Hierarchy Model by Michon (1985), these findings are of particular relevance for tasks that take place on the operational level, in which immediate reaction to (potentially dangerous) traffic situations take place in the millisecond to second range. If learning can help at making faster decisions (by triggering routines in an autonomous manner without demanding a lot of resources; Anderson, 1982), this might add to safer driving. A first consideration when implementing IDSS in cars could be to offer a training to buyers of those systems. It can be imagined that before receiving the car, drivers would have the possibility to practice driving with the unknown IDSS in either a simulator or a real car. As Experiment 3 has shown, 3 sessions with roughly 45 minutes dual-task practice were enough to significantly improve performance on both tasks and reduce dual-task costs. If training takes place in a virtual environment, learning curves based on individual driving behavior and progress could be calculated to define at which point maximum performance is reached. On a more general level, *awareness campaigns* to communicate the positive effect of practice on driving performance at more advanced age, could not only be beneficial for car vendors, but for national governments as well. Driver training at later age to overcome some deficits due to age-related changes could raise awareness among older drivers of their own driving behavior and eventually lead to increased safety on roads.

As to age-related effects on learning in dual-task driving situations, an important finding of these studies is that older adults show considerable learning progress in such a situation. Nevertheless, this finding needs to be interpreted with caution. First, although older adults benefit from practice, they never reach the same level of performance as younger adults. These findings are as expected and in accordance with literature on age-related changes, leading to slowing in cognitive processes (Borella et al., 2008; Brouwer et al., 1991; Ponds et al., 1988; Wild-Wall & Falkenstein, 2010) such as attentional processes and memory functions, which are important for skill acquisition (Anderson, 1982). Nevertheless, from a practical point of view, it is important to take into consideration that at some point, older adults will be "at their best performance" and that no further practice will improve their performance. When designing and selling IDSS, such limitations should be considered.

Another important finding concerning age-related effects on learning is that, when driving situations become rather difficult (i.e., lane-changing as compared to lane-keeping or in dual- versus single-task situations), older adults do not benefit from practice: No improvement in any of their performances took place. In other words: When the road gets rough, all resources of older adults seem to be focused on driving the car, leaving no room for concentrating on any other tasks, let alone improve them. Findings on an experiment with a relevant secondary task however, show that older adults are perfectly capable of learning in difficult dual-task situations where they are forced to divide their attention between two tasks that are (almost) equally demanding (i.e., when the secondary task is of relevance for the primary task). Within the driving context, these findings are reassuring as it shows that older adults have their priorities right (i.e., on the road) when secondary information is not fundamentally important for safe driving. However, in the case of a safety-enhancing IDSS, older adults seem to be capable of assuring a focus of attention on both the road and the secondary task.

The studies in this thesis showed as well, that learning rates differ between age groups. Older adults learn at a slower rate than younger adults. One practical implication of the findings in this thesis could be to implement a "training mode" within the IDSS, which allows gradual assistance of the IDSS until the driver feels confident enough to drive completely with it. However, this training mode must be adapted to individual needs. Younger adults would probably be much faster to learn to drive with the IDSS, in which case gradual assistance might lead to frustration. A helpful tool to foresee and analyze learning progress, is the use of learning curves. They may be used to define ceiling effects (i.e., the point in time at which practice has no measurable effect on learning anymore).

An important result and of high relevance for IDSS implementation, was the finding that a relevant secondary task (i.e., a secondary task which could not be ignored as it provided direct information for the driving task) caused more interference with the driving task than a secondary task without relevance for the driving task (i.e., a system that can be ignored as it does not provide direct information for the driving task). These interference effects were stronger for older adults as compared to younger adults. It should be noted that this effect was as expected, as visual search tasks demand discerning target item from distractor items, for which older adults need more time (Hahn et al., 2011). Taking into account the Three-Level Task Hierarchy Model by Michon (1985), a relevant system would have an effect on the operational level, that is, in the millisecond and second range, and with a high impact on driving safety. An irrelevant system could be operating on the strategic level, with relevance for the driving task (e.g., a navigation system), but without any effect on driving safety. The visual search task interfered with the visual driving task, as it forced participants to direct their attention off the road and onto the visual task display, leading to structural interference (i.e., when two tasks require the same modalities, e.g., the visual channel; Wickens, 1984). This is an important confirmation that indeed warnings or IDSS at the operational level should avoid visual representations (and employ auditory signals instead; Seppelt & Wickens, 2003; Wickens, 1984).

## 7.4 Directions for Future Research

This thesis was, to current knowledge, a first attempt at exploring the topic of learning for different age groups in dual-task driving situations. Although many useful new insights concerning this question were gathered, many issues for future research remain to be explored. A first direction comes from the artificial environment used. Although a laboratory environment allowed good experimental control, it is necessary to perform experiments like these in either high-fidelity simulators (including other traffic, different weather and traffic conditions, more realistic IDSS, etc.) or with on-road tests. More realistic driving settings have the inconvenience of reduced experimental control, but allow the potential discovery of factors not taken into account or even considered within laboratory research. Early research by Schlag (1983) has shown that several age effects observed in simulator studies disappear in real-life driving. Recent studies have confirmed as well, that older adults benefit from naturalistic settings in experimental setups (Hahn, Wild-Wall, & Falkenstein, 2013). A first indication that results in real-life driving might actually differ from driving in a simulator come from findings concerning driving with a relevant task. In the experimental setting with a relevant secondary task, older adults seemed capable of dividing attention between two tasks, whereas when the secondary task was of no relevance, they did not show this behavior. This might be an indication that in real-life situations (with different priorities?), additional factors might play a role which may compensate for cognitive effects of aging.

Another interesting perspective for future research comes from the well-known *Selection, Optimization, and Compensation* (SOC) Model for adaptive development (Baltes & Baltes, 1990). According to this model, three adaptive regulatory processes allow for compensatory behaviors or a revision of goal or task priorities when aging (Li, Krampe, & Bondar, 2005): Selection, optimization and compensation. The mechanism of *selection* is generally presented as a process of goal choice (Boker, 2013). Within the driving context it could consist of prioritizing information represented on for example a navigation display at the cost of driving performance. *Optimization* refers to the application of methods to achieve selected goals, as well as the selection of appropriate methods (Boker, 2013). For the example above, practice could help a driver to both drive and gather information on a navigation display. Finally, *compensation* is the use of alternative means when previously preferred methods become unavailable (Boker, 2013). Research has shown that especially optimization and compensation become more important for older adults between 60 and 80 years old (Freund & Baltes, 2002). Research by Li and colleagues (2005) showed that optimization and compensation strategies accounted for adaptive resource allocation, which when facing potentially competing challenges, have older adults invest most of their cognitive resources into prioritizing sensory-motor functions at the cost of cognitive performance. The outcomes of the learning experiment in which the secondary task was of relevance for the driving task however seem to indicate that when the cognitive task is more imperative, other compensatory mechanisms seem to play a role to assure a resource allocation which optimizes both the sensory-motor as well as the cognitive task. Future research could explore the mechanisms behind these findings.

Another direction for future research comes from first indications on strategy differences in learning between younger and older adults (see Experiment 3, Chapter 5). Unraveling the cognitive processes used for learning in dual-task driving situations between younger and older adults, might provide us with important insights into potential qualitative differences in learning strategies. If learning strategies truly differ, this could have important implications for the implementation and HMI-design of IDSS. One hypothesis is that younger adults are better capable of managing dual-task situations, as their executive systems allow for more parallel processing (Göthe et al., 2007). One possible way of examining the cognitive strategies involved while managing a secondary task in the driving context, is to provide older adults with a situation that challenges their control of resource allocation strategies (Li et al., 2005).

It would be interesting as well to have a closer look at retention intervals within the context of learning in dual-task driving situations for both younger and older adults. A retention interval of 3 weeks was used following the third regular practice session. Results on all performance measures remained unchanged, indicating that what had been learned was stable. This finding applied to both younger as well as older adults. Within the context of IDSS, it is probably unrealistic to not use a system for 3 weeks in a row, therefore these findings are rather reassuring, meaning that even when the IDSS is not used for a longer period of time, drivers are still able to perform at the same level when picking the system up again. The question of interest that remains though is at what point the retention interval becomes critical, especially for older adults, as they are slower to learn in a dual-task driving situation than younger adults. One direction of future research could be to focus on this "turning point" so that IDSS can be adapted accordingly (e.g., by restarting a training program if the system has not been used for a long time). Another interesting thought could be to use the retention interval as a diagnostic tool for the number of practice sessions. By varying the number of practice sessions, it can be expected that performance measures in the retention session change. This could be an indicator for the number of training sessions needed to obtain a stable learning result, unaffected by decay.

Finally it is worth considering whether learning could be used as a diagnostic tool for future limitations due to aging. All older adults in the previously presented studies were tested for cognitive fitness. However, for example in the Netherlands, older adults with mild cognitive impairments (MIC) are allowed to drive as long as they pass a driving-fitness test. The point at which an individual should cease driving is of particular concern to driving authorities, physicians and family members of older drivers who are potentially unsafe (Devlin, McGillivray, Charlton, Lowndes, & Etienne, 2012). First of all, it would be interesting to examine whether older adults with mild cognitive impairments generate the same learning results (several studies seem to predict potentially different outcomes, e.g.; Gillis, Quinn, Phillips, & Hampstead, 2013; Moulin et al., 2007). Furthermore, using learning as a diagnostic tool, it would be interesting to examine whether and to what extent learning can be used to predict age-related cognitive impairments.

## 7.5 Conclusion

The purpose of this thesis was to explore a particular question within the research domain of dual-task driving: *Can practice help older people to drive with IDSS more successfully?* The laboratory experiments presented in this thesis provide useful first insights that practice can indeed help older adults to attenuate initial age-related difficulties in a dual-task driving environment. From a practical point of view, training programs for learning to drive with IDSS may help older adults to benefit from the intended assistance. Nevertheless, although promising, caution should be taken when interpreting data from laboratory studies. Future research should focus on replicating those findings in more ecologically valid environments, to evaluate older adults' performance in more realistic situations and to examine (potential) compensation strategies (Baltes & Baltes, 1990).

# References

Ackermann, A. D. (2009). Investigation of learning outcomes for the acquisition and retention of cpr knowledge and skills learned with the use of high-fidelity simulation. *Clinical Simulation in Nursing*, *5*(6), e213-e222.

Adrian, J., Postal, V., Moessinger, M., Rascle, N., & Charles, A. (2011). Personality traits and executive functions related to on-road driving performance among older drivers. *Accident Analysis & Prevention*, *43*(5), 1652-1659.

Allain, P., Nicoleau, S., Pinon, K., Etcharry-Bouyx, F., Barré, J., Berrut, G., … Gall, D. L. (2005). Executive functioning in normal aging: A study of action planning using the zoo map test. *Brain and Cognition*, *57*(1), 4-7.

Allen, T., Lunenfeld, H., & Alexander, G. (1971). *Driver information needs* (Tech. Rep.).

Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis & Prevention*, *27*(5), 707-715.

Anderson, J. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369-403.

Anglim, J. (2011). *Strategies in skill acquisition: Reconciling continuous models of the learning curve with abrupt strategy shifts.* Unpublished doctoral dissertation.

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review*, *25*(1), 45-65.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology A*, *49*(1), 5-28.

Ball, K., Owsley, C., Stalvey, B., Roenker, D. L., Sloane, M. E., & Graves, M. (1998). Driving avoidance and functional impairment in older drivers. *Accident Analysis & Prevention*, *30*(3), 313-322.

Baltes, P., & Baltes, M. (1990). Successfull aging: Perspectives from the behavioral sciences. In P. Baltes & M. Baltes (Eds.), (p. 1-34). New York: Cambridge University Press.

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(3), 199-208.

Berthon-Donk, V., Grosjean, M., & Rinkenauer, G. (2011). Effect of feedback on performance in the lane-change test. In *Sixth international driving symposium on human factors in driver assessment, training and vehicle design* (p. 277-283).

Berthon-Donk, V., Grosjean, M., & Rinkenauer, G. (in press). Driver adaptation to information and assistance systems. In A. Stevens, C. Brusque, & J. Krems (Eds.), *Driver adaptation to information and assistance systems* (chap. Intelligent Driver Support Systems: Effects of Learning and Aging). Stevenage: The Institution Engineering and Technology.

Boker, S. (2013). Selection, optimization, compensation, and equilibrium dynamics. *The Journal of Gerontopsychology and Geriatric Psychiatry*, *26:1*, 61-73.

Borella, E., Carretti, B., & De Beni, R. (2008). Working memory and inhibition across the adult life-span. *Acta Psychologica*, *128*(1), 33-44.

Brookhuis, K. A., & de Waard, D. (1999). Limiting speed, towards an intelligent speed adapter (isa). *Transportation Research Part F: Traffic Psychology and Behaviour*, *2*(2), 81-90.

Brookhuis, K. A., De Waard, D., & Janssen, W. (2001). Behavioural impacts of advanced driver assistance systems: an overview. *Special Issue: Implementation Issues of Advanced Driver Assistance Systems, European Journal of Transport and Infrastructure Research*, *3*, 245-254.

Brouwer, W. H., Waterink, W., Van Wolffelaar, P. C., & Rothengatter, T. (1991). Divided attention in experienced young and older drivers: Lane tracking and visual analysis in a dynamic driving simulator. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *33*, 573-582.

Brouwer, W. H., Withaar, F. K., Tant, M. L. M., & van Zomeren, A. H. (2002). Attention and driving in traumatic brain injury: A question of coping with time-pressure. *The Journal of Head Trauma Rehabilitation*, *17*(1), 1-15.

Brumby, D., Salvucci, D., & Howes, A. (2007). *An empirical investigation into dual-task trade-offs while driving and dialing.* British Computer Society.

Bruyas, M.-P., Brusque, C., Auriault, A., Tattegrain, H., Aillerie, I., & Duraz, M. (2008). *Impairment of lane change performance due to distraction: Effect of experimental contexts.*

Bundesamt, S. (2009). *Bevölkerung deutschlands bis 2060 - 12. koordinierte bevölkerungsvorausberechnung.* (Tech. Rep.).

Bundesamt, S. (2011). *Verkehrsunfälle - unfälle von senioren im strassenverkehr* (Tech. Rep.). Statistisches Bundesamt.

Caird, J. K., Chisholm, S. L., & Lockhart, J. (2008). Do in-vehicle advanced signs enhance older and younger drivers intersection performance? driving simulation and eye movement results. *International Journal of Human-Computer Studies*, *66*(3), 132-144.

Carsten, O., & Brookhuis, K. A. (2005). Issues arising from the haste experiments. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*, 191-196.

Cerella, J. (1985). Information processing rates in the elderly. *Psychol Bull*, *98*(1), 67-83.

Chapman, B. P., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(2), 157-167.

Charlton, S. G., & Starkey, N. J. (2011). Driving without awareness: The effects of practice and automaticity on attention and driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(6), 456-471.

Chipman, M. L., Payne, J., & McDonough, P. (1998). To drive or not to drive: the influence of social factors on the decisions of elderly drivers. *Accid Anal Prev*, *30*(3), 299-304.

Chisholm, S. L., Caird, J. K., & Lockhart, J. (2008). The effects of practice with mp3 players on driving performance. *Accid Anal Prev*, *40*(2), 704-13.

Cooper, J., & Strayer, D. (2008). Effects of simulator practice and real-world experience on cell-phone - related driver distraction. *Human Factors*, *50*, 893-902.

Davidse, R. J. (2004). *Ouderen en its: Samen sterk(er)?* (Tech. Rep.). Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV.

Devlin, A., McGillivray, J., Charlton, J., Lowndes, G., & Etienne, V. (2012). Investigating driving behaviour of older drivers with mild cognitive impairment using a portable driving simulator. *Accident Analysis & Prevention*, *49*(0), 300-307.

Diehl, M., Willis, S. L., & Schaie, K. W. (1995). Everyday problem solving in older adults: Observational assessment and cognitive correlates. *Psychology and Aging*, *10*(3), 478-491.

Dogan, E., Steg, L., & Delhomme, P. (2011). The influence of multiple goals on driving behavior: The case of safety, time saving, and fuel saving. *Accident Analysis & Prevention*, *43*(5), 1635-1643.

Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis & Prevention*, *39*(3), 581-590.

Donmez, B., Boyle, L. N., & Lee, J. D. (2008). Mitigating driver distraction with retrospective and concurrent feedback. *Accident Analysis & Prevention*, *40*(2), 776-786.

Dotzauer, M., Berthon-Donk, V., Beggiato, M., Haupt, J., & Piccinini, G. (in press). Driver adaptation to information and assistance systems. In A. Stevens, C. Brusque, & J. Krems (Eds.), *Driver adaptation to information and assistance systems* (chap. Methods to assess behavioural adaptation over time in the context of ADAS).

Ebbinghaus, H. (1885a). *Über das gedächtnischtnis: Untersuchungen zur experimentellen pscyhologie*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Ebbinghaus, H. (1885b). *Memory: A contribution to experimental pscyhology.* New York City: Teachers College, Columbia University.

Edelman, D. A., Mattos, M. A., & Bouwman, D. L. (2010). Fls skill retention (learning) in first year surgery residents. *Journal of Surgical Research*, *163*(1), 24 - 28. doi: 10.1016/j.jss.2010.03.057

Edwards, J. D., Ross, L. A., Wadley, V. G., Clay, O. J., Crowe, M., Roenker, D. L., & Ball, K. K. (2006). The useful field of view test: Normative data for older adults. *Archives of Clinical Neuropsychology*, *21*(4), 275-286.

Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 97-120.

Fofanova, J., & Vollrath, M. (2011). Distraction while driving: The case of older drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*(0).

Freund, A., & Baltes, P. (2002). Life-management strategies of selection, optimization, and compensation: Measurement by self-report and construc validity. *Journal of Personality and Social Psychology*, *82(4)*, 642-662.

Fuller, R. (2007). *Motivational determinants of control in driving task.* Springer.

Gillis, M. M., Quinn, K. M., Phillips, P. A. T., & Hampstead, B. M. (2013). Impaired retention is responsible for temporal order memory deficits in mild cognitive impairment. *Acta Psychologica*, *143*(1), 88-95.

Gilmore, G., Royer, F., Gruhn, J., & Esson, M. (2004). Symboldigit substitution and individual differences in visual search ability. *Intelligence*, *32*, 47-64.

Gombeski, W. R., Effron, D. M., Ramirez, A. G., & Moore, T. J. (1982). Impact on retention: comparison of two cpr training programs. *American Journal of Public Health*, *72*(8), 849-852.

Gruendl, M. (2005). *Fehler und fehlverhalten als ursache von verkehrsunf$\tilde{A}$llen und konsequenzen f$\tilde{A}\frac{1}{4}$r das unfallvermeidungspotenzial und die gestaltung von fahrerassistenzsystemen.* Unpublished doctoral dissertation.

Göthe, K., Oberauer, K., & Kliegl, R. (2007). Age differences in dual-task performance after practice. *Psychology and Aging*, *22*(3), 596-606.

Hahn, M., Wild-Wall, N., & Falkenstein, M. (2011). Age-related differences in performance and stimulus processing in dual task situation. *Brain Research*, *1414*(0), 66-76.

Hahn, M., Wild-Wall, N., & Falkenstein, M. (2013). *Age-differentiated work systems* (C. Schlick, Ed.). Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-642-35057-$3_1$3

Hale, A. R., Stoop, J., & Hommels, J. (1990). Human error models as predictors of accident scenarios for designers in road transport systems. *Ergonomics*, *33*(10-11), 1377-1387.

Hancock, P. A., & Parasuraman, R. (1992). Human factors and safety in the design of intelligent vehicle-highway systems (ivhs). *Journal of Safety Research*, *23*(4), 181-198.

Harbluk, J., Burns, P., Lochner, M., & Trbovich, P. (2007). *Using the lane-change test (lct) to assess distraction: Tests of visual-manual and speech-based operation of navigation system interfaces.*

Harbluk, J., Mitroi, J., & Burns, P. (2009). Three navigation systems with three tasks: using the lane change test (lct) to assess distraction demand. In *Fifth international driving symposium on human factors in driving assessment, training and vehicle design.*

Hart, S., & Staveland, L. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human mental workload* (p. 139-183). Amsterdam: North Holland.

Heide, A., & Henning, K. (2006). The âcognitive carâ?: A roadmap for research issues in the automotive sector. *Annual Reviews in Control*, *30*(2), 197-203.

Hellier, E., Naweed, A., Walker, G., Husband, P., & Edworthy, J. (2011). The influence of auditory feedback on speed choice, violations and comfort in a driving simulation game. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(6), 591-599.

Henninger, D. E., Madden, D. J., & Huettel, S. A. (2010). Processing speed and memory mediate age-related differences in decision making. *Psychology and Aging*, *25*(2), 262-70. doi: 10.1037/a0019096

Hoedemaeker, M., & Brookhuis, K. A. (1998). Behavioural adaptation to driving with an adaptive cruise control (acc). *Transportation Research Part F: Traffic Psychology and Behaviour*, *1*(2), 95-106.

Hofmann, P. (2011). *On the influence of action preparation on steering performance in a lane change task.* Unpublished doctoral dissertation, TU Dortmund.

Horrey, W. J. (2009). On allocating the eyes: Visual attention and in-vehicle technologies. *Human Factors of Visual and Cognitive Performance in Driving, Edited by C. Castro. CRC Press, Boca Raton, Florida, USA*, 151 - 166.

Horrey, W. J., Wickens, C., & Consalus, K. (2006). Modeling driversâ visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied*, *12*(2), 67-78.

Huemer, A. K., & Vollrath, M. (2010). Alcohol-related impairment in the lane change task. *Accident Analysis & Prevention*, *42*(6), 1983-1988.

Hummel, T., Kühn, M., Bende, J., & Lang, A. (2011). *Advanced driver assistance systems: An investigation of their potential safety benefits based on an analysis of insurance claims in germany* (Tech. Rep.). German Insurance Association.

Ishihara, S. (1990). *The series of plates designed as a test for clour blindness: 24 plates edition.* Tokyo: Kanehara.

Ivnik, R., Malec, J., Smith, E., G.E.and Tangalos, & Petersen, R. (1996). Neuropsychological tests' norms above age 55: Cowat, bnt, mae token, wrat-r reading, amnart, stroop, tmt, and jlo. *The Clinical Neuropsychologist*, *10*(3), 262-278. doi: 10.1080/13854049608406689

Jackson, M. L., Croft, R. J., Kennedy, G. A., Owens, K., & Howard, M. E. (2013). Cognitive components of simulated driving performance: Sleep loss effects and predictors. *Accident Analysis & Prevention*, *50*(0), 438-444.

Jagacinski, R. J., Liao, M. J., & Fayyad, E. A. (1995). Generalized slowing in sinusoidal tracking by older adults. *Psychol Aging*, *10*(1), 8-19.

Kaczorowski, J., Levitt, C., Hammond, M., Outerbridge, E., Grad, R., Rothman, A., & Graves, L. (1998). Retention of neonatal resuscitation skills and knowledge: a randomized controlled trial. *Fam Med*, *30*(10), 705-11.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kantak, S. S., & Winstein, C. J. (2012). Learning performance distinction and memory processes for motor skills: A focused review and perspective. *Behavioural Brain Research*, *228*(1), 219-231.

Kim, J., Koubek, R., & Ritter, F. (2007). Investigation of procedural skills degradation from dfiferent modalities. In *Proceedings of the 8th international conference on cognitive modeling* (p. 255-260). Oxford, UK: Taylor & Francis/Psychology Press.

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accid Anal Prev*, *42*(3), 827-34.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol*, *77*(6), 1121-34.

Latash, M., Scholz, J., & Schöner, G. (2002). Motor control strategies revealed in the structure of motor variability. *Exercise and Sport Sciences Reviews*, *30*(1), 26-31.

Lavery, J. (1962). Retention of simple motor skills as a function of type of knowledge of results. *Canadian Journal of Psychology*, *16*, 300-311.

Levy, J., & Pashler, H. (2008). Task prioritisation in multitasking during driving: opportunity to abort a concurrent task does not insulate braking responses from dual-task slowing. *Applied Cognitive Psychology*, *22*(4), 507-525.

Lezak, M., Howieson, D., & Loring, D. (2004). *Neuropsychological assessment* (M. Lezak, Ed.). Oxford University Press.

Li, K., Krampe, R., & Bondar, A. (2005). Cognitive limitations in aging and psychopathology. In R. Engle, G. Sedek, U. Von Hecker, & D. McIntosh (Eds.), (p. 190-218). Cambridge University Press.

Llaneras, R. E., Swezey, R. W., Brock, J. F., Rogers, W. C., & Van Cott, H. P. (1998). Enhancing the safe driving performance of older commercial vehicle drivers. *International Journal of Industrial Ergonomics*, *22*(3), 217-245.

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behav Res Methods Instrum Comput*, *31*(4), 557-64.

Lundberg, C., Hakamies-Blomqvist, L., Almkvist, O., & Johansson, K. (1998). Impairments of some cognitive functions are common in crash-involved older drivers. *Accident Analysis & Prevention*, *30*(3), 371-377.

Maciej, J., & Vollrath, M. (2009). Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accident Analysis & Prevention*, *41*(5), 924-930.

Mahr, A., & Mueller, C. (2011). A schema of possible negative effects of advanced driver assistant systems. In *Sixth international driving symposium on human factors in driver assessment, training and vehicle design* (p. 116-121).

Markoff, J. (2013). *Google cars drive themselves, in traffic. www.nytimes.com* (Nos. 2013, April 02).

Mattes, S. (2003). The lane-change task as a tool for driver distraction evaluation. In H. Rausch & H. Bubb (Eds.), *Quality of work products in enterprises of the future.* Ergonomia Verlag.

Mattes, S., & Hallén, A. (2009). Surrogate distraction measurement techniques: The lane change test. In M. A. Regan, J. Lee, & K. L. Young (Eds.), *Driver distraction. theory, effects and mitigation.* (p. 107-122). Boca Raton: CRC Group Taylor & Francis Group.

McCarley, J., Vais, M., Pringle, H., Kramer, A., Irwin, D., & Strayer, D. (2004). Conversation disrupts change detection in complex traffic scenes. *Human Factors*, *46*, 424-436.

McGwin, G., Sims, R. V., Pulley, L., & Roseman, J. M. (2000). Relations among chronic medical conditions, medications, and automobile crashes in the elderly: A population-based case-control study. *American Journal of Epidemiology*, *152*(5), 424-431.

Merat, N., Anttila, V., & Luoma, J. (2005). Comparing the driving performance of average and older drivers: The effect of surrogate in-vehicle information systems. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 147-166.

Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In L. Evans & R. Schwing (Eds.), *Human behavior and traffic safety* (p. 485-520). New York: Plenum Press.

Minin, L., Benedetto, S., Pedrotti, M., Re, A., & Tesauri, F. (2012). Measuring the effects of visual demand on lateral deviation: A comparison among driverâs performance indicators. *Applied Ergonomics*, *43*(3), 486-492.

Mitsopoulos-Rubens, E., Trotter, M. J., & Lenné, M. G. (2010). Effects on driving performance of interacting with an in-vehicle music player: A comparison of three interface layout concepts for information presentation. *Applied Ergonomics*, *In Press, Corrected Proof*.

Moulin, C. J. A., Laine, M., Rinne, J. O., Kaasinen, V., Sipilä, H., Hiltunen, J., & Kangasmäki, A. (2007). Brain function during multi-trial learning in mild cognitive impairment: A pet activation study. *Brain Research*, *1136*(0), 132-141.

Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. Anderson (Ed.), *Cognitive skills and their acquisition* (p. 1-55). Hillsdale, NJ: Erlbaum.

Olive, T. (2004). Working memory in writing : Empirical evidences from the dual-task technique. *European Psychologist*, *9*, 32-42.

Owsley, C., & McGwin Jr, G. (1999). Vision impairment and driving. *Survey of Ophthalmology*, *43*(6), 535-550.

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research*, *23*(7), 689-699.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 358-377.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychol Bull*, *116*(2), 220-44.

Peretti, C.-S., Danion, J.-M., Gierski, F., & Grangé, D. (2002). Cognitive skill learning and aging: A component process analysis. *Archives of Clinical Neuropsychology*, *17*(5), 445-459.

Petzoldt, T., Bär, N., Ihle, C., & Krems, J. F. (2010). Learning effects in the the lane change task (lct) - evidence from two experimental studies. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Pohl, P. S., Winstein, C. J., & Fisher, B. E. (1996). The locus of age-related movement slowing: Sensory processing in continuous goal-directed aiming. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *51B*(2), P94-P102.

Polson, M. C., & Friedman, A. (1988). Task-sharing within and between hemispheres: a multiple-resources approach. *Hum Factors*, *30*(5), 633-43.

Ponds, W., Brouwer, W., & Wolffelaar, P. (1988). Age differences in divided attention in a simulated driving task. *Journal of Gerontology: Psychological Sciences*, *43*(6), 151-156.

Popken, A., Nilsson, A., & Krems, J. (2008). Drivers' s reliance on lane keeping assistance systems. effect of different levels of assistance. In C. Brusque (Ed.), *European conference on human interface design for intelligent transport systems* (p. 301-310). Humanist Publications.

Poston, B., Van Gemmert, A. W. A., Barduson, B., & Stelmach, G. E. (2009). Movement structure in young and elderly adults during goal-directed movements of the left and right arm. *Brain and Cognition*, *69*(1), 30-38.

Ranney, T. (2008). *Driver distraction: A review of the current state-of-knowledge* (Tech. Rep.). U.S. Department of Transportation: National Highway Traffic Safety Administration.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. In *Ieee transactions on systems, man, and cybernetics* (Vols. SMC-13, No. 3, p. 257-266).

Rasmusson, D. X., Zonderman, A. B., Kawas, C., & Resnick, S. M. (1998). Effects of age and dementia on the trail making test. *The Clinical Neuropsychologist*, *12*, 169-178.

Reitan, R., & Wolfson, R. (2003). Comprehensive handbook of psychological assessment, volume 1, intellectual and neuropsychological assessment. In M. Hersen, G. Goldstein, & S. Beers (Eds.), (chap. The Halstead-Reitan neuropsychological test battery for adults: Theoretical, methodological and validational bases). John Wiley & Sons, Inc.

Repovs, G., & Baddeley, A. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience*, *139*(1), 5-21.

Riby, L., Perfect, T., & Stollery, B. (2004). The effects of age and task domain on dual task performance: A meta-analysis. *Journal of Cognitive Psychology*, *16*(6), 863-891.

Roskam, Brookhuis, de Waard, Carsten, Read, Jamson, ... Victor (2002). *Haste deliverable 1: Development of experimental protocol* (Tech. Rep.).

Rubichi, S., Neri, M., & Nicoletti, R. (1999). Age-related slowing of control processes: Evidence from a response coordination task. *Cortex*, *35*(4), 573-582.

Rudin-Brown, C. M., & Parker, H. A. (2004). Behavioural adaptation to adaptive cruise control (acc): implications for preventive strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *7*(2), 59-76.

Russell, D. M., & Newell, K. M. (2007). On no-kr tests in motor learning, retention and transfer. *Human Movement Science*, *26*(1), 155-173.

Ruthruff, E., Hazeltine, E., & Remington, R. (2006). What causes residual dual-task interference after practice? *Psychological Research / Psychologische Forschung*, *70*, 494-503. ([1])

Salmoni, A. W., Schmidt, R., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin*, *95*(3), 355-386.

Salthouse, T. (1992). What do adult age differences in the digit symbol substitution test reflect? *Journal of gerontology: Psychological Sciences*, *47*, 121-128.

Salvucci, D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, *48*(2), 362-380.

Schermer, F. (1998). *Lernen und gedächtnis (2. überarbeitete und erweiterte auflage).* Stuttgart: Verlag W. Kohlhammer.

Schlag, B. (1983). Elderly drivers in germany–fitness and driving behavior. *Accident Analysis & Prevention*, *25*, 47-55.

Schmidt, R. A., Lange, C., & Young, D. E. (1990). Optimizing summary knowledge of results for skill learning. *Human Movement Science*, *9*(3-5), 325-348.

Seppelt, B., & Wickens, C. (2003). *In-vehicle tasks: Effects of modality, driving relevance, and redundancy* (Tech. Rep.). Warren, MI: General Motors Corporation.

Shanmugaratnam, S., Kass, S. J., & Arruda, J. E. (2010). Age differences in cognitive and psychomotor abilities and simulated driving. *Accident Analysis & Prevention*, *42*(3), 802-808.

Shinar, D., Tractinsky, N., & Compton, R. (2005). Effects of practice, age, and task demands, on interference from a phone task while driving. *Accid Anal Prev*, *37*(2), 315-26.

Sülzenbrück, S., & Heuer, H. (2011). Type of visual feedback during practice influences the precision of the acquired internal model of a complex visuo-motor transformation. *Ergonomics*, *54*(1), 34-46. doi: http://dx.doi.org/10.1080/00140139.2010.535023

Smiley, A. (1999). Adaptive strategies of older drivers. In 36-43 (Ed.), *Conference of transportation in an aging society: A decade of experience* (Vol. 27). Transportation Research Board of the National Academies.

Sorel, O., & Pennequin, V. (2008). Aging of the planning process: The role of executive functioning. *Brain and Cognition*, *66*, 196-201.

Speelman, C. (2005). Beyond the learning curve: The construction of mind. In C. Speelman & K. Kirsner (Eds.), (p. 26-64). Oxford Scholarship Online. doi: 10.1093/acprof:oso/9780198570417.003.0002

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *12*(6), 462-466.

Summala, H. (1997). Hierarchical model of behavioural adaptation and traffic accidents. In T. Rothengatter & E. Carbonell Vaya (Eds.), *Traffic and transportation psychology* (p. 41-52). Amsterdam: Pergamon.

Sweeney, J. A., Rosano, C., Berman, R. A., & Luna, B. (2001). Inhibitory control of attention declines more than working memory during normal aging. *Neurobiology of Aging*, *22*(1), 39-47.

Tattegrain, H., Bruyas, M.-P., & Karmann, B. (2009). Comparison between adaptive and basic model metrics in lane change test to assess in-vehicle secondary task demand. In *21st esv conference.*

Totzke, I., Hofmann, M., Meilinger, T., Rauch, N., & Schmidt, G. (2004). *Kompetenzerwerb für fahrerinformationssysteme: Einfluss des lernprozesse auf die interaktion mit fahrerassistenzssystemen.* Offenbach: Berthold Druck.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97-136.

Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics*, *50*(8), 1235-1249.

Van Engelen, D. (2011). *Attention drivers! - analyzing driver distraction.* Unpublished doctoral dissertation.

Verwey, W. B., Abrahamse, E., Ruitenberg, M., Jiménez, L., & Kleine, d. E. (2011). Motor skill learning in the middle-aged: limited development of motor chunks and explicit sequence knowledge. *Psychological Research*, *75*, 406-422.

Victor, T. W., Harbluk, J. L., & Engström, J. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 167-190.

Voelcker-Rehage, C., & Alberts, J. (2007). Effect of motor practice on dual-task performance in older adults. *Journal of Gerontology: Psychological Sciences*, *62B*, 141-148.

Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, *10*(4), 293-309.

Vollrath, M., Schleicher, S., & Gelau, C. (2011). The influence of cruise control and adaptive cruise control on driving behaviour  a driving simulator study. *Accident Analysis & Prevention*, *43*(3), 1134-1139.

Wagner, A. (2006). Encoding and retrieval from long-term memory. In E. E. Smith & S. M. Kosslyn (Eds.), *Cognitive psychology: Mind and brain.* Prentice Hall.

Wechsler, D. (1944). *The measurement of adult intelligence* (Third ed.). Baltimore: The Williams & Wilkins Company.

Wickens, C. (1984). Processing resources in attention. In R. Parasuraman & D. Davies (Eds.), *Varieties of attention* (p. 239-257). Hillsdale, NJ: Erlbaum.

Wickens, C. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomical Science*, *3*(2), 159-177.

Wickens, C. (2007). Time sharing and multiple resource theory. In C. Wickens & J. S. McCarley (Eds.), *Applied attention theory.* New York: CRC Press.

Wickens, C. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 449-455.

Wickens, C., & Seppelt, B. (2002). *Interference with driving or in-vehicle task information: The effects of auditory versus visual delivery* (Tech. Rep. No. AHFD-02-18/GM-02-3). Warren, MI: General Motors Corporation.

Wilde, G. J. S. (1989). Accident countermeasures and behavioural compensation: The position of risk homeostasis theory. *Journal of Occupational Accidents*, *10*(4), 267-292.

Wild-Wall, N., & Falkenstein, M. (2010). Age-dependent impairment of auditory processing under spatially focused and divided attention: An electrophysiological study. *Biological Psychology*, *83*(1), 27-36.

Wilschut, E. (2009). *The impact of in-vehicle information systems on simulated driving performance.* Unpublished doctoral dissertation.

Wilschut, E., Rinkenauer, G., Brookhuis, K. A., & Falkenstein, M. (2008). *Effects of visual search task complexity on lane change task performance.* HUMANIST Publications.

Wolfe, J. M. (1994). Visual search in continuous, naturalistic stimuli. *Vision Research*, *34*(9), 1187-1195.

Young, K. L., Lenné, M. G., & Williamson, A. R. (2011). Sensitivity of the lane change test as a measure of in-vehicle system demand. *Appl Ergon*, *42*(4), 611-8.

Zhao, G., & Wu, C. (2012). The effects of driver identity on driving safety in a retrospective feedback system. *Accident Analysis & Prevention*, *45*(0), 354-365.

Zijlstra, F., & Van Doorn, T. (1985). *The construction of a scale to measure perceived effort.* Unpublished doctoral dissertation.

# List of Figures

1.1 Roadmap representing the development of IDSS over time and their increase in complexity (this figure was adapted from Heide & Henning, 2006). 4
1.2 Wickens' four-dimensional multiple resource model (this figure was reproduced by permission of Van Engelen, 2011, Figure 3.2, Page 19). . . . . . 8
1.3 Percentage of accidents on German roads in 2011 in which an older driver was at fault, as a function of driving errors (this figure was adapted from Statistisches Bundesamt, 2011, Figure 6, p. 11). . . . . . . . . . . . . . . 11
1.4 Examples of driving tasks combining Michon's Three-Level Task Hierarchy Model and Rasmussen's Skill-Knowledge-Rule Model (this figure was adapted from Hale et al.; 1990, Figure 1, p. 1383). . . . . . . . . . . . . . 22

2.1 Screenshot of the LCT showing the straight 3-lane road without other traffic and with lane-change signs at both sides of the road which, in this case, indicate a lane change to the left of 1 lane. . . . . . . . . . . . . . . 31
2.2 Experimental setting representing the position of the LCT projection, the seating position of the participant, the position of the VST display and the function of different computers used (see text for details). . . . . . . 32
2.3 Example of the visual search task used in the HASTE project. The encircled symbols represent a target. . . . . . . . . . . . . . . . . . . . . . . . 33
2.4 VST examples with (a) and without presence of a target (b). . . . . . . . 34
2.5 Timely apparition of the LCT sign and the VST displays in the dual-task condition. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36

2.6 Example of the electrooculogram (EOG) PC display. The first 4 channels record eye movements (vertical [v] EOG upper electrode, vEOG lower electrode, horizontal [h] EOG left electrode, hEOG right electrode), the channel XCO and 10 represent information provided by the LCT PC (the participants current position on the track, the optimal track) and the channel Diode_Scree represents a channel with information from the Eprime PC (the diode lighted up when the VST appeared and turned off as soon as a response was detected). Finally there are some events represented, provided by the Eprime PC as well, concerning precise position of the LCT sign (LC-Sign), the moment a button was pressed (VST-$\text{Response}_{\text{Onset}}$) and some VST information concerning the type of VST sent (type of target, target present/absent, set size). . . . . . . . . . . . . 37
2.7 Code table used in the Digit Symbol Substitution Test with digit-pairs (above) and an array of random digits from 1 to 9 (below) with space below where the participant has to add the correct symbol according to the code table. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40
2.8 Representation of a lane-change maneuver (upper figure) and the resulting numerical derivation (lower figure) based on which the beginning ($\text{LC}_{\text{Onset}}$) and the end ($\text{LC}_{\text{Offset}}$) of a lane-change movement are calculated. The x-axes of both figures represent the time on the trial in milliseconds (ms). The y-axis of the upper figure represents the lateral position in meters (m). Zero (0) represents the middle of the middle lane. The y-axis of the lower figure represents the lateral vehicle velocity in meters per second (m/s). The red dashed lines in both figures represent the moments in time at which lateral vehicle velocity reached 20% before and after the peak lateral vehicle velocity. . . . . . . . . . . . . . . . . . . . . . . . . . 45
2.9 Representation of different points in time within one trial and how they combine to account for relative lane-change and lane-keeping segments, as well as independent temporal variables for the LCT, the VST and the dual-task condition (see text for details). . . . . . . . . . . . . . . . . . . 47

2.10 Illustration on how lateral deviation measures are obtained: The green line represents the reference trajectory. The red line represents the driving behavior of the participant and his deviation from the reference trajectory (orange area). LCT signs (black vertical dashed lines) are indicated at their actual position. The eyes as well as the grey vertical dashed lines represent visibility of the LCT signs (LC-Sign$_{\text{Onset}}$ at 40 m before the actual position of the LCT sign. The green vertical dashed lines illustrate the moment at which the reference trajectory starts the lane change (at 30 m before the actual position of the LCT sign). . . . . . . . . . . . . . 48
2.11 Vertical saccade between the LCT display and the VST display. . . . . . 51

3.1 Examples of different types of lane-change maneuvers divided into Lane Change (LC) and Lane Keeping (LK) segments either by using absolute windows (a and c) or by using relative windows (b and d). The x-axis represents time, the y-axis represents the lateral position of the vehicle. The green line represents the reference trajectory. The red line represents the driving behavior of a slow participant, the blue line represents the driving behavior of a fast participant. The orange area represents the deviation from the reference trajectory. The vertical dashed lines illustrate the beginning and the end of LC and LK segments. Note that these figures serve as examples and are not to scale. . . . . . . . . . . . . . . . 55
3.2 Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and age group (younger, older). . . . . . . . . 61
3.3 Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping) and window (absolute, relative). . . . . . . . . . . . . . . . 62
3.4 Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping). . . 62
3.5 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . . 63
3.6 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of (a) task type (single-LCT, dual task) and window (absolute, relative). . . 65

3.7 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single, dual), window (absolute, relative), and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . 65
3.8 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . 66
3.9 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . 68

4.1 Mean Movement Time (MT-LC) as a function of feedback condition (feedback, no feedback) and age group (younger, older). . . . . . . . . . . . . 83
4.2 Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . . . . . . . . . . 84
4.3 Mean Lateral Deviation (MDEV) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 85
4.4 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . 86
4.5 Mean Standard Deviation of Steering Wheel angle (SDSW) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping). . . . . . . . . . . . . . . . 86
4.6 Panel (a): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for older adults. Panel (b): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for younger adults. . . . . . . . . . . . . . . . . . 88
4.7 Mean NASA-TLX ratings for Temporal Demand as a function of task type (single-LCT, single-VST, dual-task) and feedback condition (feedback, ne feedback). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89

4.8 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . 91
4.9 Mean Lateral Deviation (MDEV) as a function of block (1-10) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . . . . 92
4.10 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . . . . . . 93
4.11 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), feedback condition (feedback, no feedback) and age group (young, old). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93

5.1 Mean Lateral Deviation (MDEV) as a function of (a) session (1-4) and segment (lane change, lane keeping) and (b) as a function of session (1-4) and task type (single-LCT, dual-task). . . . . . . . . . . . . . . . . . . . 109
5.2 Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping). 111
5.3 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4) and task type (single-LCT, dual). . . . . . . . . . . . . . . . 112
5.4 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 113
5.5 Proportion of misses as a function of session (1-4), segment (lane change, lane keeping) and age group (younger, older). . . . . . . . . . . . . . . . 115
5.6 Proportion of misses as a function of session (1-4), age group (younger, older) and set size (9, 16, 25). . . . . . . . . . . . . . . . . . . . . . . . 116
5.7 Mean Rating Scale of Mental Effort (RSME) as a function of session (1-4) and task type (single-LCT, single-VST, DUAL). . . . . . . . . . . . . . . 117
5.8 Mean vertical saccade frequency (VSF) as a function of session (1-4), age group (younger, older) and segment (lane change, lane keeping). . . . . . 118
5.9 Total Glance Duration (TGD) as a function of session (1-4) and age group (young, old). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 119
5.10 Mean Movement Time (MT-LC) as a function of session (1-4), age group (younger, older) and set size (9, 16, 25). . . . . . . . . . . . . . . . . . . 120

5.11 Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older). . . . 121
5.12 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older). 122
5.13 Proportion of incorrect responses as a function of session (1-4), age group (young, old) and set size (9, 16, 25). . . . . . . . . . . . . . . . . . . . . . 123

6.1 Examples of an (a) relevant the Visual Search Task (VST) and b non-relevant VST. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 137
6.2 Mean Inter-Response Interval (IRI) as a function of dual-task type (dual, dual-r) and set size (9, 25). . . . . . . . . . . . . . . . . . . . . . . . . 145
6.3 Mean Inter-Response Interval (IRI) as a function of (a) dual-task type (dual, dual-r) and age (younger, older). . . . . . . . . . . . . . . . . . . . 146
6.4 Mean Lateral Deviation (MDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25). . . . . . . . . . . . . . . 147
6.5 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25). . . 148
6.6 Proportion of LCT-Misses as a function of session (1-4) and dual-task type (dual, dual-r) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions. . . . . . . . . . . . . . . 150
6.7 Proportion of LCT-Errors (a) session (1-4), age (younger, older) and dual-task type (dual, dual-r) for (a) non-arcsine transformed and (b) arcsine-transformed values, with fitted power functions for the older age group. . 151
6.8 Mean Reaction Time until Lane Change (RT-LC) as a function of session (1-4), dual-task type (dual, dual-r) and set size (9, 25) for (a) older adults and (b) younger adults. . . . . . . . . . . . . . . . . . . . . . . . . . . 152

A.1 Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping) and age group (younger, older) . . . . . . . . . . . . . . . . ii
A.2 Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and age group (younger, older). . . . . . . . . iii
A.3 Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping) and window (absolute, relative). . . . . . . . . . . . . . . . iv

A.4 Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping). . . iv
A.5 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . . v
A.6 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of (a) task type (single-LCT, dual task) and window (absolute, relative) and (b) task type (single-LCT, dual task) and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . vii
A.7 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single, dual), window (absolute, relative), and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . viii
A.8 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . ix
A.9 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . x

B.1 Mean Reaction Time until Lane Change as a function of task type (single, dual) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . xiii
B.2 Mean Movement Time (MT-LC) as a function of (a) task type (single, dual) and age group (younger, older) and (b) as a function of feedback condition (feedback, no feedback) and age group (younger, older). . . . . xiii
B.3 Mean Lateral Deviation (MDEV) as a function of task type (single, dual) and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . xv
B.4 Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . . . . . . . . . . . xv
B.5 Mean Lateral Deviation (MDEV) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xvi
B.6 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single, dual) and segment (lane change, lane keeping). . . . . . xvii

B.7 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . xviii
B.8 Mean Standard Deviation of Steering Wheel angle (SDSW) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping). . . . . . . . . . . . . . . . xix
B.9 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . xx
B.10 Proportion of incorrect responses as a function of task type (single, dual) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . xxiii
B.11 Mean NASA-TLX ratings for Mental Demand as a function of task type (single-LCT, single-VST, dual-task) and age group (younger, older). . . . xxv
B.12 Mean NASA-TLX ratings for Temporal Demand as a function of task type (single-LCT, single-VST, dual-task) and feedback condition (feedback, ne feedback). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xxv
B.13 Mean NASA-TLX ratings for Performance as a function of task type (single-LCT, single-VST, dual-task), feedback condition (feedback, no feedback) and age group (younger, older). . . . . . . . . . . . . . . . . . xxvi
B.14 Mean Lateral Deviation (MDEV) as a function of block (1-10) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . . . xxviii
B.15 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback). . . . . . . . . . . . . . . . . . . . . . . xxviii
B.16 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), feedback condition (feedback, no feedback) and age group (young, old). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xxix
B.17 Mean Reaction Time on the Visual Search Task (RT-VST) as a function of block (1-10) and age group (young, old). . . . . . . . . . . . . . . . . . xxx

B.18 Panel (a): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for older adults. Panel (b): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for younger adults. . . . . . . . . . . . . . . . . . . xxxii

C.1 Reaction Time until Lane Change (a) and Movement Time (b) as a function of session (1-4). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xxxiv
C.2 Mean Lateral Deviation (MDEV) as a function of (a) session (1-4) and segment (lane change, lane keeping) and (b) as a function of session (1-4) and task type (single-LCT, dual-task). . . . . . . . . . . . . . . . . . . . . xxxv
C.3 Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping). xxxvii
C.4 Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older). . . . xxxviii
C.5 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older).xxxviii
C.6 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single-LCT, dual) and segment (lane change, lane keeping). . . xl
C.7 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4) and task type (single-LCT, dual). . . . . . . . . . . . . . . . xli
C.8 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xlii
C.9 Reaction Time on the Visual Search Task (RT-VST) as a function of session (1-4). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xliii
C.10 Proportion of incorrect responses as a function of session (1-4). . . . . . . xliv
C.11 Proportion of incorrect responses as a function of set size (9, 16, 25) and age group (young, old). . . . . . . . . . . . . . . . . . . . . . . . . . . . xlv
C.12 Proportion of incorrect responses as a function of session (1-4), age group (young, old) and set size (9, 16, 25). . . . . . . . . . . . . . . . . . . . . . xlvi
C.13 Mean Rating Scale of Mental Effort (RSME) as a function of (a) session (1-4) and task type (single-LCT, single-VST, DUAL) and (b) task type (single-LCT, single-VST, DUAL) and age group (younger, older). . . . . xlvii

C.14 Mean Movement Time (MT-LC) as a function of session (1-4), age group (younger, older) and set size (9, 16, 25). . . . . . . . . . . . . . . . . . . . l
C.15 Standard Deviation of Lateral Deviation (SDDEV) as a function of set size (9, 16, 25) and segment (lane change, lane keeping). . . . . . . . . . li
C.16 Mean Reaction Time on the Visual Search Task (RT-VST) as a function of session (1-4) and segment (lane change, lane keeping). . . . . . . . . . liii
C.17 Proportion of misses as a function of session (1-4), segment (lane change, lane keeping) and age group (younger, older). . . . . . . . . . . . . . . . . liv
C.18 Proportion of misses as a function of session (1-4), age group (younger, older) and set size (9, 16, 25). . . . . . . . . . . . . . . . . . . . . . . . . . lv
C.19 Inter-Response Interval (IRI) as a function of session (1-4). . . . . . . . . lvii
C.20 Mean vertical saccade frequency (VSF) as a function of session (1-4), age group (younger, older) and segment (lane change, lane keeping). . . . . . lviii
C.21 Mean Vertical Saccade Frequency (VSF) as a function of set size (9, 16, 25) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . lix
C.22 Mean Glance Duration (MGD) as a function of set size (9, 16, 25) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . lx
C.23 Mean Glance Duration (MGD) as a function of set size (9, 16, 25) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . lx
C.24 Total Glance Duration (TGD) as a function of (a) session (1-4) and age group (young, old) and as a function of (b) set size (9, 16, 25) and age group (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . lxii

D.1 Mean Reaction Time until Lane Change (RT-LC) as a function of session (1-4) and age (younger, older) with fitted power function for the older age group. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . lxiv
D.2 Mean Movement Time (MT-LC) as a function of task type (single-LCT, dual, dual-r) and age (younger, older). . . . . . . . . . . . . . . . . . . . lxv
D.3 Mean Lateral Deviation (MDEV) as a function of session (1-4) and age (younger, older) with fitted power function for the older age group. . . . lxvi
D.4 Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of session (1-4). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . lxviii

D.5 Mean Reaction Time on the Visual Search Task (RT-VST) as a function of session (1-4), age (younger, older) and set size (9, 25) with fitted power functions for all conditions. . . . . . . . . . . . . . . . . . . . . . . . . . . lxix
D.6 Mean Proportion of Incorrect Responses on the Visual Search Task as a function of task type (single-VST, dual, dual-r) and age (younger, older). lxxi
D.7 Proportion of Incorrect Responses as a function of session (1-4) and age (younger, older) for (a) non-arcsine transformed and (b) arcsine-transformed values with fitted power functions. . . . . . . . . . . . . . . lxxii
D.8 Proportion of Incorrect Responses as a function of session (1-4), age (younger, older) and set size (9, 25) for (a) non-arcsine transformed values and (b) arcsine-transformed values. . . . . . . . . . . . . . . . . . . . . . . lxxiii
D.9 Mean Proportion of Incorrect Responses on the Visual Search Task as a function of session (1-4) and task type (single-VST, dual, dual-r) with fitted power functions for the older adults, set size 9 and 25 and the younger adults, set size 25. . . . . . . . . . . . . . . . . . . . . . . . . . . lxxiii
D.10 Proportion of misses on the Visual Search Task as a function of session (1-4), age (younger, older) and task type (single-VST, dual, dual-r) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions. . . . . . . . . . . . . . . . . . . . . . . . . . . lxxiv
D.11 Proportion of misses on the Visual Search Task as a function of task type (single-VST, dual, dual-r), age (younger, older) and set size (9, 25). . . . lxxvi
D.12 Proportion of misses on the Visual Search Task as a function of session (1-4), age (younger, older) and set size (9, 25) for (a) non-arcsine transformed values and (b) arcsine-transformed values, with fitted power functions for the younger age group. . . . . . . . . . . . . . . . . . . . . . . . . . . lxxvii
D.13 Mean Rating Scale of Mental Effort (RSME) as a function of session (1-4), age (younger, older) and task type (single-LCT, single-VST, dual, dual-r), with fitted power functions for the older adults in the single-VST, the dual and the dual-r condition. . . . . . . . . . . . . . . . . . . . . . . . . lxxix
D.14 Mean NASA-TLX scores for Mental Demand as a function of (a) session (1-4) and age (younger, older), with fitted power function for the younger adults, and (b) task type (single-LCT, single-VST, dual, dual-r) and age (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . lxxxi

D.15 Mean NASA-TLX scores for Temporal Demand as a function of (a) session (1-4) and task type (single-LCT, single-VST, dual, dual-r), with fitted power functions for the single-VST, dual and dual-r conditions. . . . . . . lxxxii
D.16 Mean NASA-TLX scores for Performance as a function of session (1-4) and task type (single-LCT, single-VST, dual, dual-r). . . . . . . . . . . . lxxxiv
D.17 Proportion of LCT-Errors (a) session (1-4), age (younger, older) and dual-task type (dual, dual-r) for (a) non-arcsine transformed and (b) arcsine-transformed values, with fitted power functions for the older age group. . lxxxvii
D.18 Proportion of LCT-Errors as a function of set size (9, 25) and age (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . lxxxvii
D.19 Proportion of LCT-Misses as a function of session (1-4) and age (younger, older) for (a) non-arcsine transformed values and (b) arcsine-transformed values, with fitted power function for the older adults. . . . . . . . . . . . lxxxix
D.20 Proportion of LCT-Misses as a function of session (1-4) and set size (9, 25) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions. . . . . . . . . . . . . . . . . . . . . . . . lxxxix
D.21 Proportion of LCT-Misses as a function of session (1-4) and dual-task type (dual, dual-r) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions. . . . . . . . . . . . . . . xc
D.22 Mean Reaction Time until Lane Change (RT-LC) as a function of session (1-4), dual-task type (dual, dual-r) and set size (9, 25) for (a) older adults and (b) younger adults. . . . . . . . . . . . . . . . . . . . . . . . . . . . . xcii
D.23 Mean Lateral Deviation (MDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25). . . . . . . . . . . . . . . xcii
D.24 Mean Lateral Deviation (MDEV) as a function of session (1-4), age (younger, older) and set size (9, 25). . . . . . . . . . . . . . . . . . . . . . . . . xciii
D.25 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25). . . xciv
D.26 Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4), age (younger, older) and set size (9, 25). . . . . . . . . . . . xcv
D.27 Mean Inter-Response Interval (IRI) as a function of (a) dual-task type (dual, dual-r) and age (younger, older) and (b) dual-task type (dual, dual-r) and set size (9, 25). . . . . . . . . . . . . . . . . . . . . . . . . . . xcvi

D.28 Mean Inter-Response Interval (IRI) as a function of (a) session (1-4) and dual-task type (dual, dual-r), with fitted power function for dual-r and (b) session (1-4) and set size (9, 25), with fitted power function for set size 9. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xcviii

D.29 Mean Inter-Response Interval (IRI) as a function of set size (9, 25) and age (younger, older). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xcix

D.30 Mean Inter-Response Interval (IRI) as a function of session (1-4), age (younger, older) and set size (9, 25). . . . . . . . . . . . . . . . . . . . . . . . xcix

# Selbstständigkeitserklärung

Hiermit versichere ich, dass die vorliegende Arbeit von mir persönlich und eigenständig verfasst wurde und ich keine anderen als die im Text angegebenen Quellen und Hilfsmittel verwendet habe.

Dortmund, im Juli 2013

Vera Donk

# A Appendix 1: Results Experiment 1

In this results section, we compare results on LCT performance measures as a function of absolute and relative calculation windows by comparing single- versus dual-task data. Note that due to the amount of data, figures were only included for significant or close to significant ($p$ =.08) main effects or interactions. Note that close to significant interactions will only be described, without providing any exploratory follow-up tests.

## A.1 Comparing Single versus Dual-Task Data

Separate 4-way mixed-factors ANOVAs were conducted on the driving measures MDEV, SDDEV and SDSW. The ANOVA included the between-participants factor *age* (younger, older) to examine the effect of age on the driving measures, the within-participants factor *task type* (single, dual) to examine the effect of task complexity, the within-participants factor *segment* (LC, LK) to take into account the effect of driving difficulty on those measures and finally the within-participants factor *window* (absolute, relative) to examine the effect of window calculation method on those driving measures.

*MDEV.* The ANOVA revealed a significant main effect of task type on MDEV ($F$[1, 21] = 8.95, $\eta_p^2$ = .30, $p$ < .01) with MDEV being significantly higher in dual-task conditions (M = .82 m) as compared to single-LCT conditions (M = .77 m). There was furthermore a main effect of window ($F$[1, 21] = 220.18, $\eta_p^2$ = .91, $p$ < .001): MDEVs calculated with relative windows were higher (M = .88 m) than those calculated with absolute windows (M = .71 m). The factor segment showed a main effect as well ($F$[1, 21] = 271.55, $\eta_p^2$ = .93, $p$ < .001). Although trivial, MDEVs in LC segments were higher (M = 1.16 m) as compared to MDEVs in LK segments (M = .43 m). Finally there was a significant main effect of age ($F$[1, 21] = 36.71, $\eta_p^2$ = .63, $p$ < .001). Older adults yielded significantly higher MDEV values (M = .98 m) than younger adults (M = .61 m).

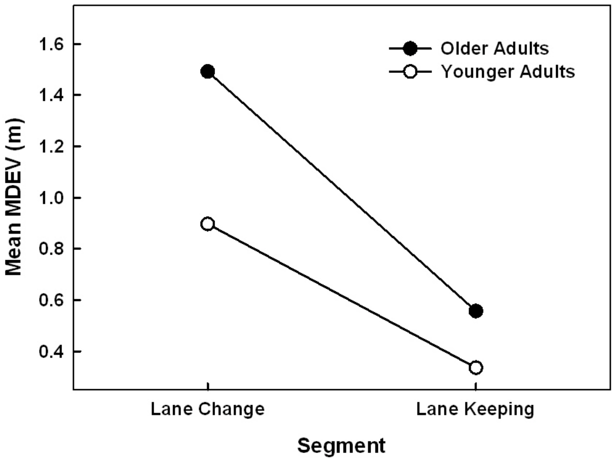


Figure A.1: Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping) and age group (younger, older)

Those main effects were however mediated by significant 2-way interactions. Figure A.1 presents mean MDEV as a function of segment and age group. There was a significant interaction between those factors ($F[1,\ 21] = 17.04$, $\eta_p^2 = .45$, $p < .001$). To follow up on this interaction we conducted two separate 1-way (segment) ANOVAs for each age group respectively. There was a significant simple main effect of the factor segment on both younger adults ($F[1,\ 10] = 265.66$, $\eta_p^2 = .96$, $p < .001$) as well as older adults ($F[1,\ 10] = 141.79$, $\eta_p^2 = .93$, $p < .001$). The interaction comes from the fact that the decrease in MDEV is stronger for older adults as compared to younger adults going from the LC segment to the LK segment.

Figure A.2 presents mean MDEV as a function of task type, window and age group. There was a significant 2-way interaction between task type and age ($F[1,\ 21] = 5.05$, $\eta_p^2 = .19$, $p < .05$) as well as task type and window ($F[1,\ 21] = 12.91$, $\eta_p^2 = .38$, $p < .01$). These interactions were however modulated by a significant 3-way (age x task type x window) interaction ($F[1,\ 21] = 5.38$, $\eta_p^2 = .20$, $p < .05$). To follow up on this interaction we performed two separate 2-way (task type x window) ANOVAs for each age group respectively. There was no signficant interaction between those two factors for the younger adults ($p > .10$), but the interaction reached significance for the older adults

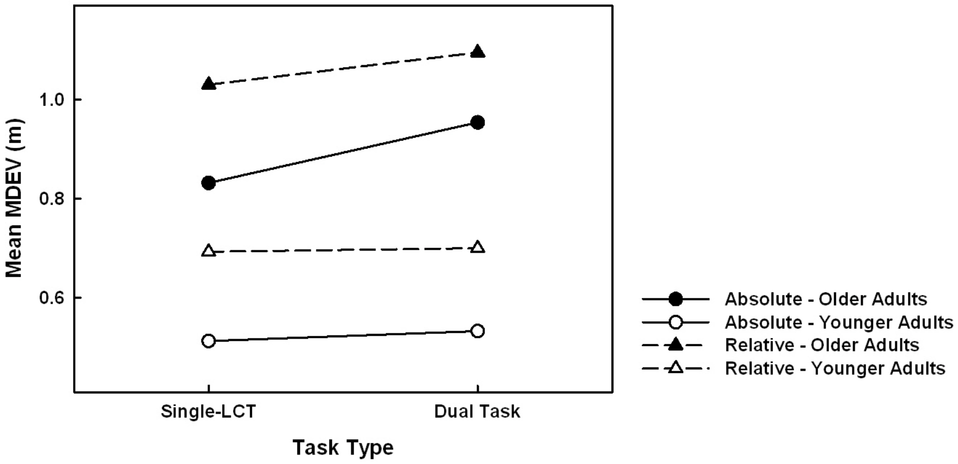


Figure A.2: Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and age group (younger, older).

($F[1,\ 11] = 10.76$, $\eta_p^2 = .49$, $p < .01$). A follow-up on the 2-way interaction for older adults with two separate 1-way (task type) ANOVAs for absolute and relative windows respectively showed that there was a significant simple main effect of task type when using the constant window ($F[1,\ 11] = 13.30$, $\eta_p^2 = .55$, $p < .01$), but, although showing a tendency, not for the relative window ($p = .07$). This means that for older adults, MDEV values increased significantly in dual-task conditions as compared to single-LCT conditions when using the absolute window for calculating MDEV. This might indeed indicate that using an absolute window for calculating MDEV values might lead to an over-estimation of those values for older adults.

Figure A.3 presents mean MDEV as a function of segment and window. The 2-way interaction between those factors was significant ($F[1,\ 21] = 120.64$, $\eta_p^2 = .85$, $p < .001$). A follow-up on this interaction with two separate 1-way (segment) ANOVAs for absolute and relative windows respectively, showed that there was a significant simple main effect of the factor segment on both absolute ($F[1,\ 22] = 102.28$, $\eta_p^2 = .82$, $p < .001$) as well as relative windows ($F[1,\ 22] = 211.98$, $\eta_p^2 = .91$, $p < .001$). The interaction comes from the fact that the decrease in MDEV is stronger going from the LC segment to the LK segment when using relative windows as compared to absolute windows for

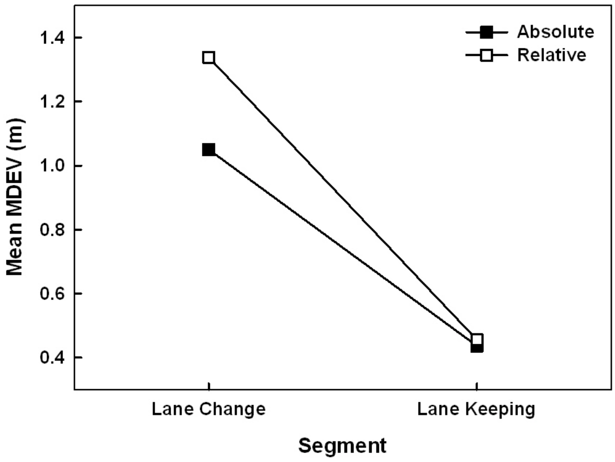


Figure A.3: Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping) and window (absolute, relative).

the calculation of MDEV.

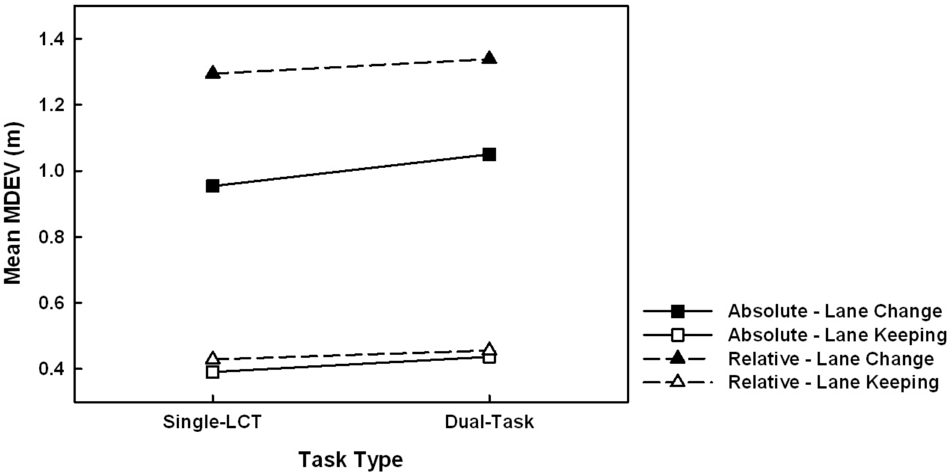


Figure A.4: Mean Lateral Deviation (MDEV) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping).

Figure A.4 shows mean MDEV as a function of task type, window and segment. The 3-way interation including those factors showed a tendency, but just missed to reach significance ($F[1,\ 21] = 3.54$, $\eta_p^2 = .14$, $p = .07$). It seems to provide however an indication that MDEV values in LK segments hardly differed independent of task type and window used for calculation. However, for the LC segments, not only did MDEV values increase in the dual-task condition as compared to the single-LCT condition, but this effect was especially noticeable when using the absolute windows to calculate MDEV. No other significant interactions with the factor task type could be found (all $p$s $> .35$).

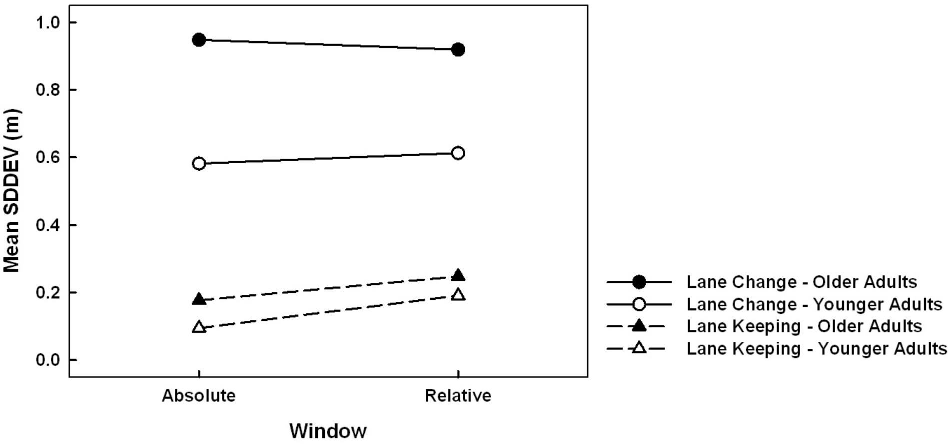
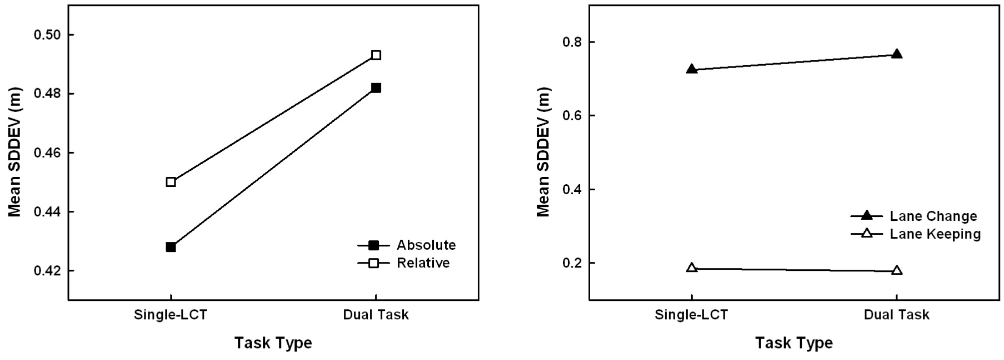


Figure A.5: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older).

*SDDEV.* Figure A.5 presents mean SDDEV as a function of window, segment and age group. The ANOVA revealed a main significant effect of window on SDDEV ($F[1,\ 21] = 99.57$, $\eta_p^2 = .83$ $p < .001$): SDDEV in relative windows was higher (M = .49 m) than SDDEV in absolute windows (M = .44 m). There was furthermore a main effect of segment ($F[1,\ 21] = 321.17$, $\eta_p^2 = .94$, $p < .001$) with lane keeping variability being higher in LC (M = .75 m) as compared to LK segments (M = .18 m). Finally there was a significant main effect of age as well ($F[1,\ 21] = 35.07$, $\eta_p^2 = .63$, $p < .001$). Older adults showed more variability (M = .56 m) around the optimal track than younger adults (M = .37

m). All those main effects were however modulated by significant 2-way interactions. The ANOVA revealed a main significant 2-way interaction between the factors window and age on SDDEV ($F[1, 21] = 25.41$, $\eta_p^2 = .55$, $p < .001$) as well as segment and age ($F[1, 21] = 16.51$, $\eta_p^2 = .44$, $p < .01$) and finally window and segment ($F[1, 21] = 89.63$, $\eta_p^2 = .81$, $p < .001$). All those 2-way interactions were modulated by a significant 3-way (window x segment x age) interaction ($F[1, 21] = 6.46$, $\eta_p^2 = .24$, $p < .05$). To follow up on this interaction, we conducted two separate 2-way (window x age) ANOVAs for LC and LK segments respectively. The interaction turned out to be significant for both LC ($F[1, 21] = 24.96$, $\eta_p^2 = .54$, $p < .001$) as well as LK segments ($F[1, 21] = 4.59$, $\eta_p^2 = .18$, $p < .05$). A further analysis with four separate 1-way (window) ANOVAs for each age group in each driving segment, revealed, for LC segments, a significant simple main effect of window for younger adults ($F[1, 10] = 21.66$, $\eta_p^2 = .68$, $p < .01$) as well as older adults ($F[1, 10] = 5.74$, $\eta_p^2 = .34$, $p < .05$). For LK segments, there was a significant simple main effect of window for both younger ($F[1, 10] = 357.01$, $\eta_p^2 = .97$, $p < .001$) and older adults as well ($F[1, 11] = 100.58$, $\eta_p^2 = .90$, $p < .001$). The 3-way interaction shown in Figure A.5 can thus be explained as follows: in LC segments, SDDEV values of older adults decrease when the relative calculation method is used, whereas for younger adults, SDDEV values increase with the relative calculation method. In LK segments however, SDDEV values for both age groups increase when the relative calculation method is used, but this increase is stronger for younger adults than for older adults.

Figure A.6a presents mean SDDEV as a function of task type and window and Figure A.6b presents mean SDDEV as a function of task type and segment. Task type interacted significantly with the factor window ($F[1, 21] = 9.11$, $\eta_p^2 = .30$, $p < .01$). To follow up on this interaction we conducted two separate 1-way (task type) ANOVAs for each window respectively. The ANOVA showed that there was no main effect of task type when using the relative window as a basis for calculating SDDEV ($p > .26$), and that this effect only showed a tendency when using the absolute window as a basis for calculation ($F[1, 22] = 3.76$, $\eta_p^2 = .15$, $p = .07$). This indicates that the 2-way interaction comes from the fact that the effect for window is bigger for single-LCT as compared to dual-task conditions. The factor task type furthermore interacted significantly with the factor segment ($F[1, 21] = 5.97$, $\eta_p^2 = .22$, $p < .05$). To follow up on this ANOVA we conducted two separate 1-way (task type) ANOVAs for each segment



(a) (b)

Figure A.6: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of (a) task type (single-LCT, dual task) and window (absolute, relative) and (b) task type (single-LCT, dual task) and segment (lane change, lane keeping).

respectively. There was a signficant simple main effect of task type in LC segments ($F$[1, 22] = 4.34, $\eta_p^2 = .17$, $p < .05$), but not for LK segments ($p = .09$). This indicates that task type had no effect in LK segments, but that in the LC segments, SDDEV increased significantly in the dual-task condition as compared to the single-LCT condition.

Figure A.7 presents mean SDDEV as a function of task type, window and segment. The 3-way interaction between those factors just missed to reach significance ($F$[1, 21] = 4.28, $\eta_p^2 = .17$, $p = .05$) but nevertheless provides an indication that although the use of relative windows yielded a bit higher SDDEV values than the use of absolute windows in LK segments, values hardly differed between single-LCT and dual-task conditions. However, for the LC segments, SDDEV values increased going from the single- to the dual-task driving condition and this effect seemed to be a bit stronger for calculations in which the absolute window was used.

*SDSW.* Figure A.8 presents mean SDSW as a function of window, segment and age group. The ANOVA revealed a significant main effect of window on SDSW ($F$[1, 21] = 135.50, $\eta_p^2 = .87$, $p < .001$): SDSW was significantly higher when the relative window calculation method was used (M = 8.56°) as compared to the use of the absolute window

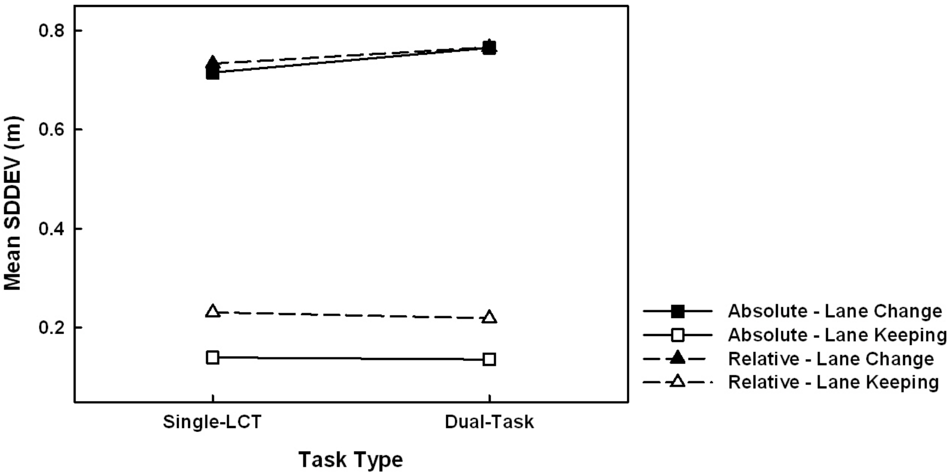


Figure A.7: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single, dual), window (absolute, relative), and segment (lane change, lane keeping).

calculation (M = 5.99°). There was a significant main effect of segment on SDSW values as well ($F$[1, 21] = 254.29, $\eta_p^2$ = .92, $p < .001$) with SDSW values being higher in LC (M = 12.89°) as compared to LK segments (M = 1.67°). Finally, a significant main effect of age on SDSW was observed ($F$[1, 21] = 30.66, $\eta_p^2$ = .59, $p < .001$): SDSW of younger adults was higher (M = 9.30°) than SDSW of older adults (M = 5.26°). Those main effects were however modulated by significant 2-way interactions. First, the ANOVA revealed a significant interaction between the factors window and age on SDSW ($F$[1, 21] = 36.09, $\eta_p^2$ = .69, $p < .001$). There was furthermore a significant 2-way interaction between segment and age ($F$[1, 21] = 49.02, $\eta_p^2$ = .70, $p < .001$) and finally the 2-way interaction between the factors window and segment turned out to be significant ($F$[1, 21] = 78.97, $\eta_p^2$ = .79, $p < .001$). All these 2-way interactions were however modulated by a significant 3-way interaction including all factors (window x segment x age) ($F$[1, 21] = 47.24, $\eta_p^2$ = .69, $p < .001$). To follow-up on this interaction we conducted two separate 2-way (window x age) ANOVAs for LC and LK respectively. The 2-way interaction was significant for both LC ($F$[1, 21] = 43.89, $\eta_p^2$ = .68, $p < .001$) as well as LK segments ($F$[1, 21] = 6.76, $\eta_p^2$ = .24, $p < .05$). To further analyse these interactions, we looked into detail at LC as well as LK segments with 4 separate 1-way (window) ANOVAs for each age group and each segment. As to the LC segments, there was a significant simple main effect of window for older adults ($F$[1, 11] = 12.55, $\eta_p^2$ = .53, $p$ < .01) as well as younger adults ($F$[1, 10] = 137.46, $\eta_p^2$ = .93, $p$ < .001). For the LK segments, a similar pattern was found: a significant simple main effect of window for older ($F$[1, 11] = 36.39, $\eta_p^2$ = .77, $p$ < .001) as well as younger adults ($F$[1, 10] = 119.39, $\eta_p^2$ = .92, $p$ < .001). The 3-way interaction shown in Figure A.8 can thus be explained as follows: in LC segments, SDSW values increase when relative calculation methods are used and this increase is strongest for younger adults. For LK segments, SDSW values are higher when the relative calcuation method is used as compared to the absolute calculation method, but again, this increase is strongest for younger adults.

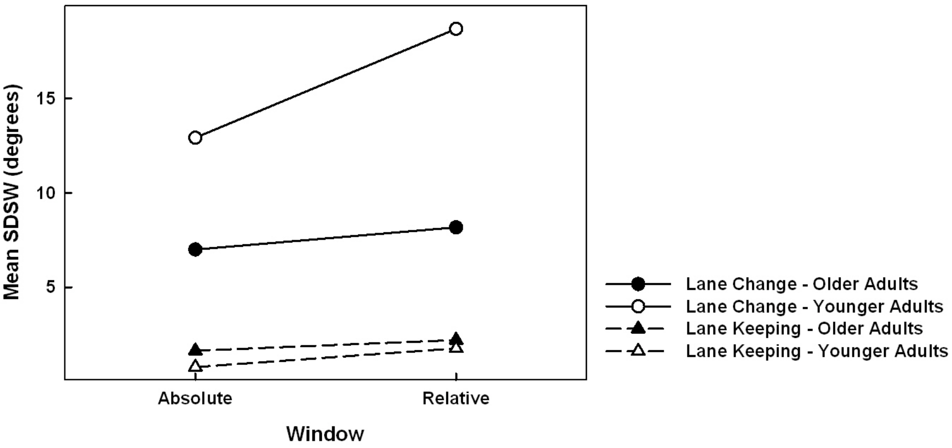


Figure A.8: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of window (absolute, relative), segment (lane change, lane keeping) and age group (younger, older).

Figure A.9 presents mean SDSW as a function of task type, window and segment. The ANOVA yielded a main significant effect of task type on SDSW ($F$[1, 21] = 46.32, $\eta_p^2$ = .69, $p$ < .001): steering-wheel variability was higher in single-LCT conditions (M = 7.92) as compared to dual-task conditions (M = 6.64). The factor task type interacted significantly with the factor window as well ($F$[1, 21] = 55.84, $\eta_p^2$ = .73, $p$ < .001) and there

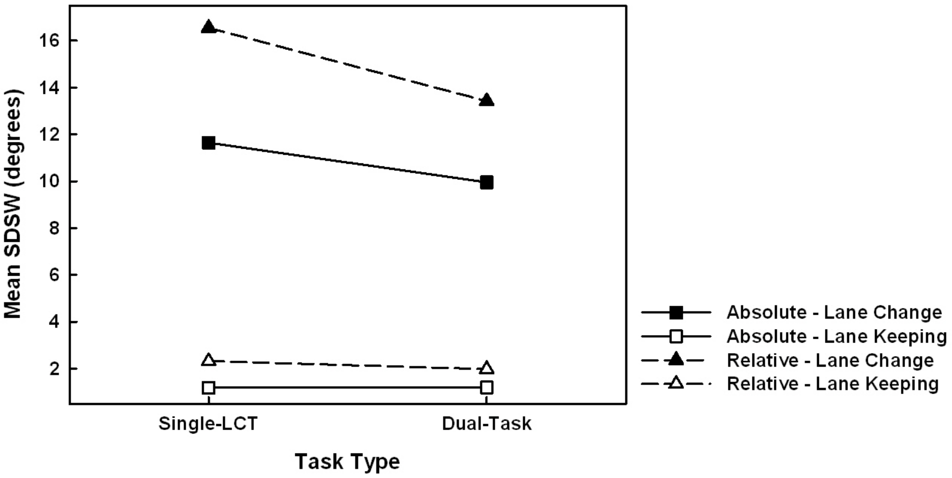


Figure A.9: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of task type (single, dual), window (absolute, relative) and segment (lane change, lane keeping).

was a significant 2-way interaction including the factors task type and segment ($F$[1, 21] = 45.57, $\eta_p^2$ = .69, $p < .001$). These 2-way interactions were however modulated by a significant 3-way (task type x window x segment) interaction ($F$[1, 21] = 25.57, $\eta_p^2$ = .55, $p < .001$). To follow up on this interaction we conducted two separate 2-way (task type x window) interactions for LC and LK segments respectively. The interaction was significant for LC ($F$[1, 22] = 40.15, $\eta_p^2$ = .65, $p < .001$) as well as LK segments ($F$[1, 22] = 35.08, $\eta_p^2$ = .62, $p < .001$). The observed 3-way interaction in Figure A.9 can thus be explained by the fact that both 2-way interactions differ according to the segment which is analyzed. A follow-up analysis of the LK segment with two separate 1-way (task type) ANOVAs for each window respectively, revealed that there was no significant simple main effect of task type for absolute windows ($p > .73$), whereas there was a significant simple main effect of task type for relative windows ($F$[1, 22] = 8.39, $\eta_p^2$ = .28, $p < .01$). The two-way interaction between task type and window for LK segments can thus be explained by the fact that there is no difference in SDSW between single-LCT and dual-task conditions when the absolute window calculation is used, whereas when the relative window calculation is used, SDSW is significantly higher in the single-LCT condition as compared to the dual-task condition. A follow-up analysis of the LC segment with two separate 1-way (task type) ANOVAs for each window respectively, showed that there was a significant simple main effect of task type for both absolute ($F[1, 22] = 44.77$, $\eta_p^2 = .67$, $p < .001$) as well as relative windows ($F[1, 22] = 48.00$, $\eta_p^2 = .69$, $p < .001$). The 2-way interaction between task type and window for LC segments can thus be explained by the fact that the difference between the absolute and the relative window calculation is bigger for the single-LCT condition as compared to the dual-task condition.

# B Appendix 2: Results Experiment 2

## B.1 Comparing Single versus Dual-Task Data

### B.1.1 Driving Measures

Separate 3-way mixed-factors ANOVAs were conducted on the driving measures RT-LC and MT-LC. The ANOVA included the within-participants factor *task type* (single, dual) to examine differences in performance between the single- and the dual-task condition, the between-participants factor *age* to examine the effect of age of the participants (*age*: younger, older) and whether participants had been submitted to the feedback or no feedback condition with the factor *feedback*. As described in the data-analysis section of the general method (Chapter **??**), RT-LC and MT-LC measures were based on LC segments only. The driving measures MDEV, SDDEV, SDSW underwent separate 4-way ANOVAs including the between-participants factors *age* and *feedback* and the within-participants factors *task type* and *segment* (LC, LK).

*RT-LC.* Figure B.1 presents mean RT-LC as a function of task type and age group. There was a significant main effect of task type ($F$[1, 46] = 53.14, $\eta_p^2$ =.54, $p$ <.001) on RT-LC. Initiation of a lane change was longer in the dual-task condition (M = 920 ms) as compared to the single-task condition (M = 859 ms). This might indicate an interference effect from the VST on the driving task. There was furthermore a significant main effect of age ($F$[1, 46] = 39.73, $\eta_p^2$ =.46, $p$ <.001): older participants needed more time to start a lane change (M = 996 ms) as compared to the younger participants (M = 784 ms). The interaction between age and task type just missed significance ($F$[1, 46] = 3.23, $\eta_p^2$ =.07, $p$ =.08) but may be an indication that older adults suffer a bit more from the addition of a secondary task to the driving task as compared to younger adults. No other main effects or interactions attained significance (all $p$s >.34).

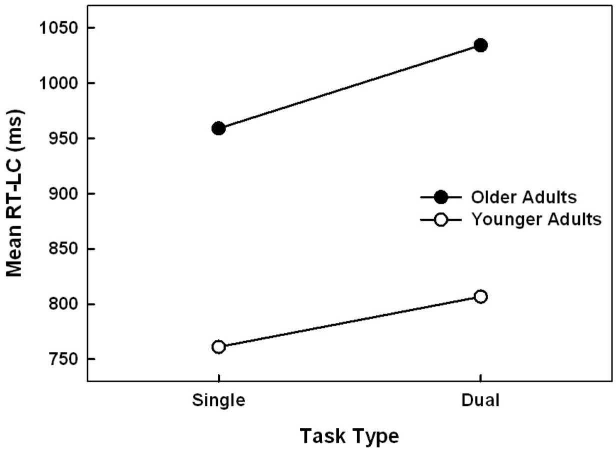
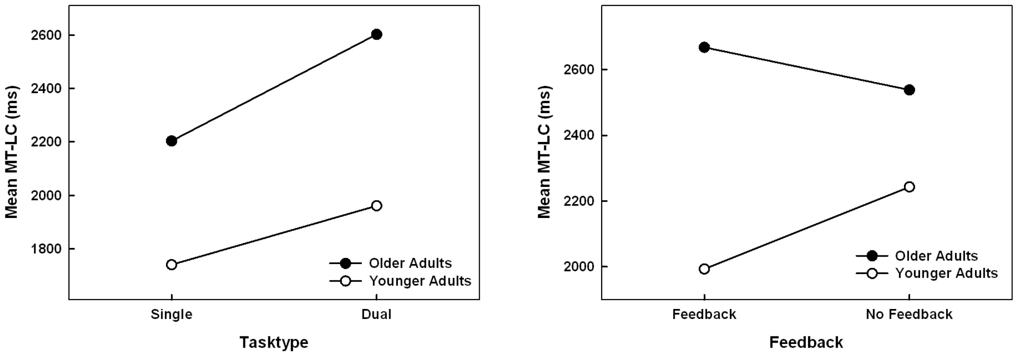


Figure B.1: Mean Reaction Time until Lane Change as a function of task type (single, dual) and age group (younger, older).



(a) (b)

Figure B.2: Mean Movement Time (MT-LC) as a function of (a) task type (single, dual) and age group (younger, older) and (b) as a function of feedback condition (feedback, no feedback) and age group (younger, older).

*MT-LC*. Figure B.2a presents mean MT-LC as a function of task type and age group. The ANOVA revealed a significant main effect of task type ($F$[1, 46] = 94.22, $\eta_p^2$ =.67, $p$ <.001) on MT-LC with movement times in the dual-task condition being significantly higher (M = 2281 ms) than those in the single-task condition (M = 1972 ms). The ANOVA revealed furthermore that movement times of older participants were significantly higher (M = 2403 ms) than that of younger participants (M = 1851 ms) as well ($F$[1, 46] = 48.98, $\eta_p^2$ =.52, $p$ <.001). The observed interaction of task type with the factor age was significant ($F$[1, 46] = 7.77, $\eta_p^2$ =.15, $p < .01$). A post-hoc analysis with paired-samples $t$-tests showed that the increase in movement time for both younger ($t$[23] = 5.71, $p < .001$) and older adults ($t$[25] = 7.76, $p < .001$) was significant going from the easier single-task condition to the more demanding dual-task condition, but the resulting increase in MT-LC was stronger for older than for younger adults. Figure B.2b presents MT-LC as a function of feedback condition and age group. A significant interaction between the factors age and feedback condition can be observed ($F$[1, 46] = 5.39, $\eta_p^2$ =.11, $p < .05$): An independant-samples $t$-test for the older adults showed that their movement time did not differ between feedback conditions ($p > .13$), whereas for younger adults, movement times significantly decreased when feedback was provided as compared to the condition without feedback ($t$[22] = 2.57, $p < .01$, one-tailed). This might be an indication that with feedback, younger adults are more concentrated on the driving task and as a result MT-LC decreases. All other main effects or interactions did not reach significance (all $p$s > .27).

*MDEV*. Figure B.3 presents mean MDEV as a function of task type and segment. There was a significant main effect of task type ($F$[1, 46] = 61.20, $\eta_p^2$ =.57, $p < .001$): MDEV in dual-task situations was significantly higher (M = 1.10 m) than in single-task situations (M = 1.04 m). There was furthermore a trivial significant main effect of segment as well ($F$[1, 46] = 3086.70, $\eta_p^2$ =.99, $p$ =.001) with MDEV in LC segments being significantly higher (M = 1.74 m) than MDEV in LK segments (M =.41 m). As can be seen in the graph, task type interacts significantly with the factor segment as well ($F$[1, 46] = 37.86, $\eta_p^2$ =.45, $p < .001$). A follow-up with paired-samples $t$-tests showed that the factor task type had no effect in LK segments ($p > .12$, one-tailed), but for LC segments, MDEV increased going from the single- to the dual-task condition ($t$[49] = 8.05, $p < .001$, one-tailed).

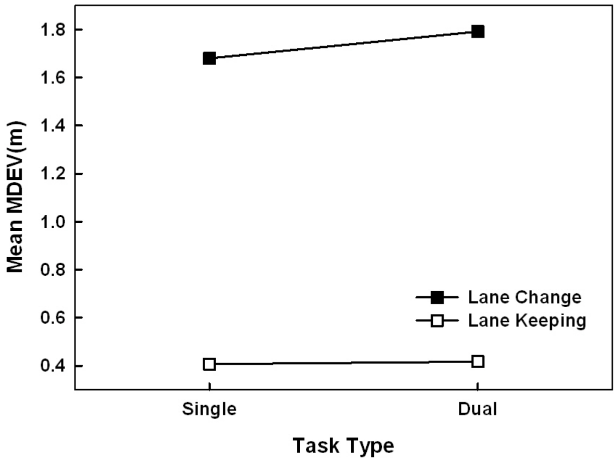


Figure B.3: Mean Lateral Deviation (MDEV) as a function of task type (single, dual) and segment (lane change, lane keeping).

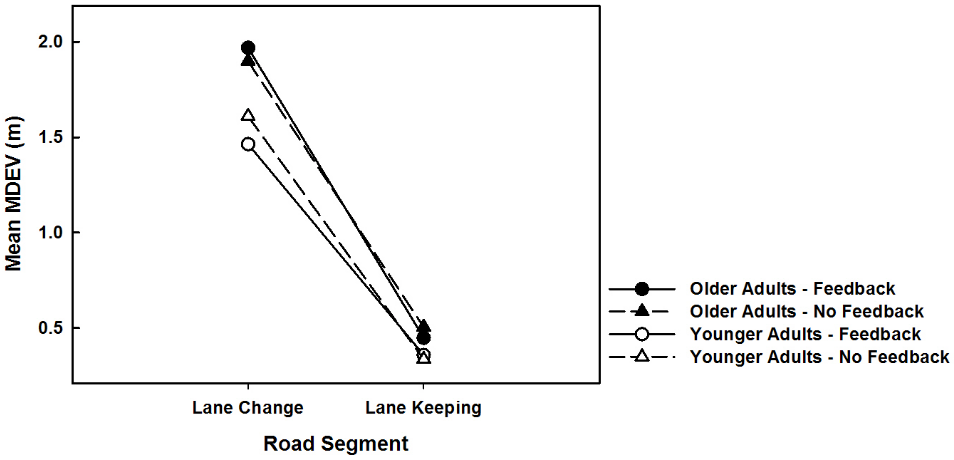


Figure B.4: Mean Lateral Deviation (MDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback).

Figure B.4 presents MDEV as a function of segment, age group and feedback condition. As can be seen in the graph, there was as significant main effect of age ($F$[1, 46] = 64.87, $\eta_p^2 = .59$, $p < .001$) which shows that MDEV of older adults was higher (M = 1.21 m) than that of younger adults (M = .94 m). The interaction between the factors segment and age was significant ($F$[1, 46] = 31.66, $\eta_p^2 = .41$, $p < .001$), but modulated by a significant 3-way (age x segment x feedback) interaction ($F$[1, 46] = 9.56, $\eta_p^2 = .17$, $p < .01$). A post-hoc analysis with two separate 2-way (segment x feedback) ANOVAs for younger and older adults respectively, showed a significant 2-way interaction between segment and feedback condition ($F$[1, 22] = 5.10, $\eta_p^2 = .19$, $p < .05$) for younger adults: independant samples $t$-tests showed that values in the LC segments differed significantly ($t$[22] = 1.81, $p < .05$, one-tailed), whereas values values for the LK segments did not differ statistically ($p > .26$, one-tailed). For older adults, the 2-way interaction between segment and feedback was significant as well ($F$[1, 22] = 4.36, $\eta_p^2 = .15$, $p < .05$): independant samples $t$-tests however showed that there was no difference between LC values ($p > .15$, one-tailed) with and without feedback as well as LK values ($p > .06$, one-tailed) with and without feedback.

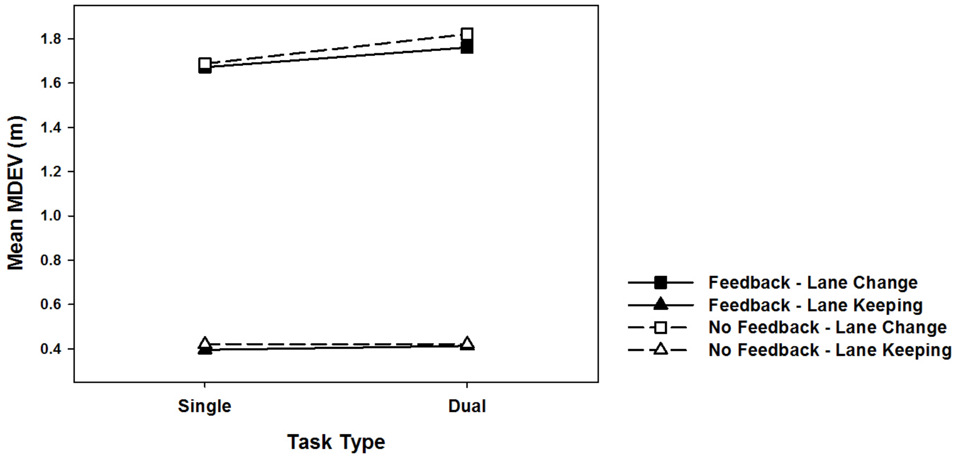


Figure B.5: Mean Lateral Deviation (MDEV) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping).

Figure B.5 presents mean MDEV as a function of task type, segment and feedback condition. The 3-way interaction including those three factors showed a tendency but just missed to reach significance ($F[1, 46] = 3.33$, $\eta_p^2 = .07$, $p = .07$). It seems to indicate however that although feedback had hardly any effect in LK segments, it seems to play a role in LC segments. When feedback is provided, the increase in MDEV is less going from the single- to the dual-task condition than when no-feedback is provided. This might indicate a possible positive effect of feedback on priority management (i.e. priority on the driving task) when the driving task is rendered more difficult by a distracting secondary task.

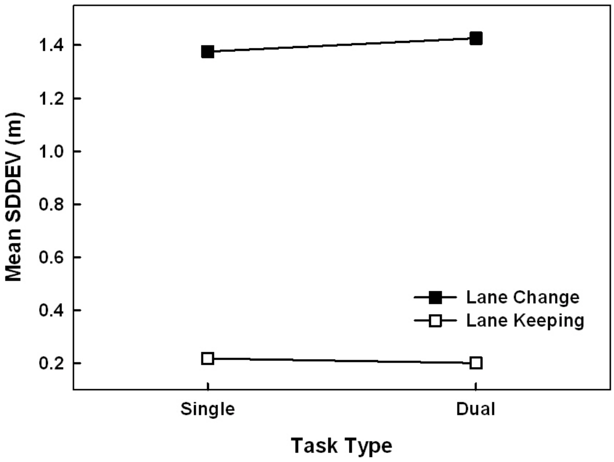


Figure B.6: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single, dual) and segment (lane change, lane keeping).

*SDDEV.* Figure B.6 presents mean SDDEV as a function of task type and segment. The factor task type showed a significant main effect on SDDEV ($F[1, 46] = 8.21$, $\eta_p^2 = .15$, $p < .01$) and variability around the optimal track was significantly higher in the dual-task condition (M = .81 m) as compared to the single-task condition (M = .80 m). There was a significant main effect of segment as well ($F[1, 46] = 4688.09$, $\eta_p^2 = .99$, $p < .001$): SDDEV in LC segments was higher (M = 1.40 m) than that of LK segments (M = .21 m), indicating more variability around the optimal track in the LC segments. The observed interaction between task type and segment was significant ($F[1, 46] =$

27.18, $\eta_p^2 = .45$, $p < .001$). Paired-samples $t$-tests showed that observed values in both LC segments ($t[49] = 4.60$, $p < .001$, one-tailed) as well as LK segments ($t[49] = 3.13$, $p < .01$, one-tailed) differed between single- and dual-task conditions. The significant interaction comes from the fact that in LC segments SDDEV increases in the dual-task condition, whereas in the LK segments SDDEV decreases in the dual-task condition.

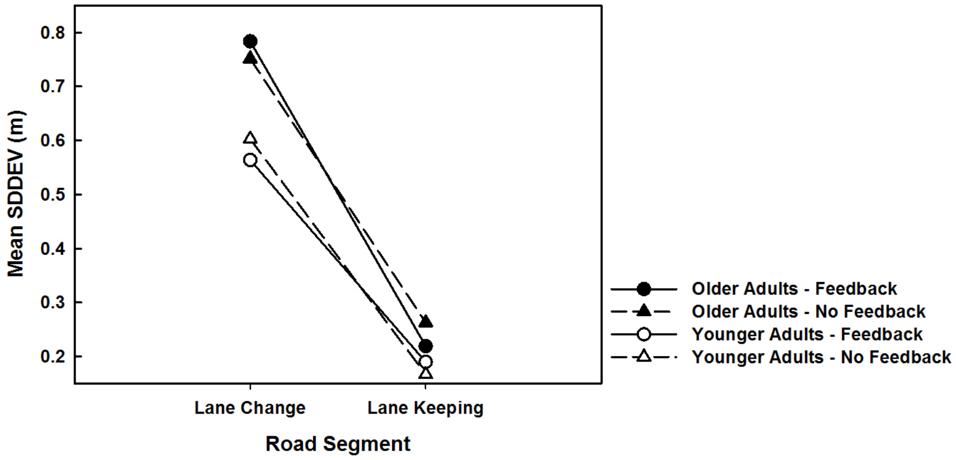


Figure B.7: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback).

Figure B.7 presents SDDEV as a function of segment , feedback condition and age group. There was a significant main effect of age ($F[1, 46] = 55.15$, $\eta_p^2 = .55$, $p < .001$) with SDDEV of older adults being significantly higher (M = .87 m) than that of younger adults (M = .74 m). There was furthermore a significant 2-way interaction between the factors segment and age ($F[1, 46] = 18.56$, $\eta_p^2 = .29$, $p < .001$). This interaction was however modulated by a significant 3-way (segment x feedback x age) interaction ($F[1, 46] = 8.19$, $\eta_p^2 = .15$, $p < .01$). To further analyze this interaction we conducted two separate 2-way (segment x feedback) ANOVAs for the older and the younger adults respectively. The 2-way interaction between segment and feedback condition turned out to be significant for younger adults ($F[1, 22] = 4.61$, $\eta_p^2 = .17$, $p < .05$), but, despite a tendency, missed to reach significance for older adults ($p = .08$). Independent-samples

$t$-tests on the data of younger adults showed however that SDDEV, despite feedback, only showed a strong tendency in LC ($p$ = .05), but no significant effect in LK segments ($p$ = .10) from the condition without feedback. The interaction shown in Figure B.7 can thus be explained by the fact that feedback had a strong effect on lane-change variability for younger adults in LC segments only. No other significant main effects or interactions were found for the omnibus ANOVA (all $p$s > .13).

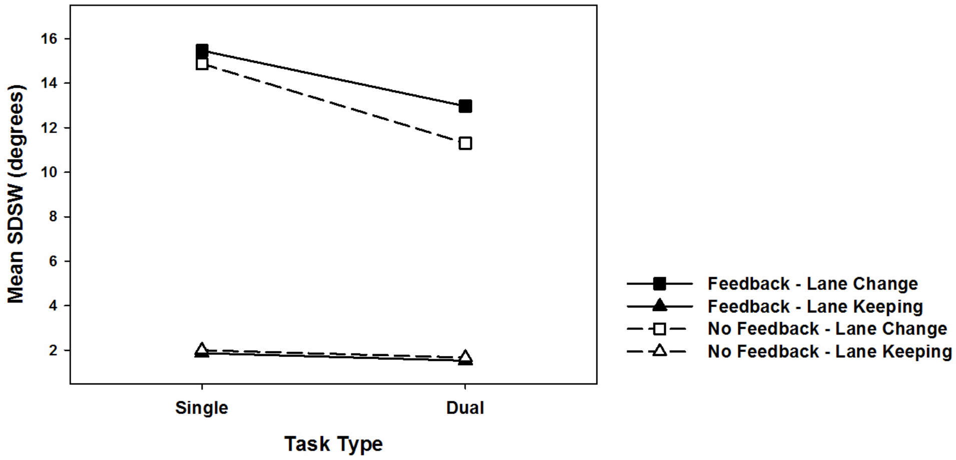


Figure B.8: Mean Standard Deviation of Steering Wheel angle (SDSW) as a function of task type (single, dual), feedback condition (feedback, no feedback) and road segment (lane change, lane keeping).

*SDSW.* Figure B.8 presents SDSW as a function of task type, segment and feedback condition. The ANOVA revealed a significant main effect of task type on SDSW ($F$[1, 46] = 107.29, $\eta_p^2$ = .70, $p$ < .001): SDSW was higher in the single-task condition (M = 8.55°) as compared to the dual-task condition (M = 6.87°). The ANOVA revealed a significant main effect of segment as well ($F$[1, 46] = 880.86, $\eta_p^2$ = .95, $p$ < .001) with SDSW being significantly higher in the LC segments (M = 13.66°) as compared to the LK segments (M = 1.77°). The factor task type interacted significantly with the factor segment ($F$[1, 46] = 106.25, $\eta_p^2$ = .70, $p$ < .001), but this interaction was modulated by a significant 3-way (task type x segment x feedback condition) interaction ($F$[1, 46] = 4.32, $\eta_p^2$ = .09, $p$ < .05). To analyze this 3-way interaction in detail we conducted two separate 2-way (task type x feedback condition) ANOVAs for LC and LK segments respectively. The 2-way interaction between task type and feedback did not reach significance for the LK segments ($p > .91$). For the LC segments, this interaction showed a tendency, but just missed to reach significance ($F[1, 48] = 3.49$, $\eta_p^2 = .07$, $p = .07$). This indicates that SDSW has a tendency to decrease in the dual-task condition as compared to the single-task condition especially when no feedback was provided.

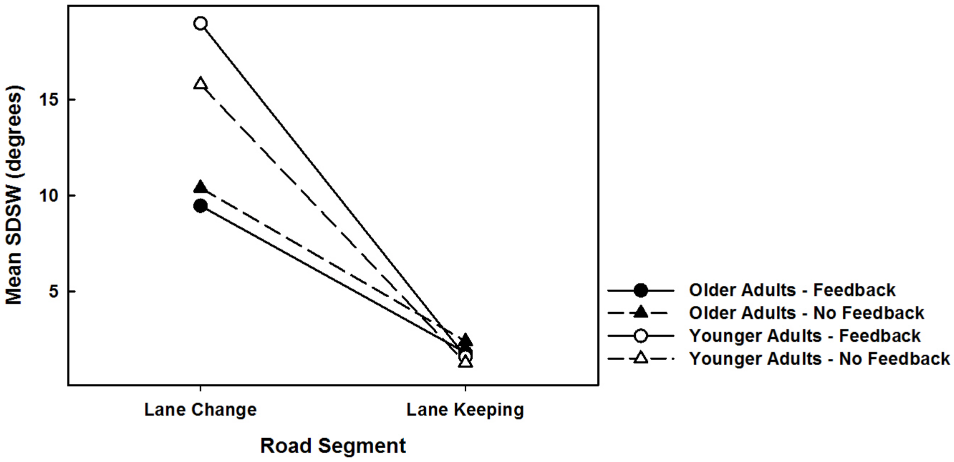


Figure B.9: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of segment (lane change, lane keeping), age group (younger, older) and feedback condition (feedback, no feedback).

Figure B.9 presents SDSW as a function of segment, age group and feedback condition. There was a significant main effect of age ($F[1, 46] = 63.10$, $\eta_p^2 = .58$, $p < .001$): SDSW of younger adults was higher (M = 9.41°) than SDSW of older adults (M = 6.01°). The observed interaction between the factors segment and age was significant ($F[1, 46] = 102.64$, $\eta_p^2 = .69$, $p < .001$). Independent-samples $t$-tests showed however that there was no difference between the FB- and NOFB-condition for LC ($p > .23$) and LK segments ($p > .20$). The observed 2-way interaction between age and feedback condition was significant as well ($F[1, 46] = 8.65$, $\eta_p^2 = .16$, $p < .01$). Independent-samples $t$-tests showed that SDSW values differed significantly for younger adults between the FB- and the NOFB-condition ($t[22] = 2.51$, $p < .05$, one-tailed). The $t$-test showed however no significant difference between SDSW values in the FB- and the NOFB-condition for older adults ($p$ = .07). The 3-way interaction between the factors segment, age and feedback showed a tendency, but just missed to reach significance ($F[1, 46] = 3.99$, $\eta_p^2 = .08$, $p = .05$). It strongly indicates however that younger adults' SDSW values were higher in LC segments, but especially when feedback was provided. Older adults' SDSW values were higher in LC segments as well, but feedback did not have an effect on those values. No other significant main effects or interactions were observed (all $p$s $> .11$).

### B.1.2 Visual Search Measures

Performance on the VST was analyzed by the RT-VST as well as computing the proportion of incorrect responses for each condition and participant. As no misses occurred in the VST single-task condition, no ANOVA for the measure misses was performed. Values for incorrect responses were arcsine transformed to deal with the non-normality of proportions (e.g., see Winer, 1971). To establish the influence of task difficulty, feedback condition and age on both RT-VST and the proportion of incorrect responses, separate 3-way mixed-factors ANOVAs were performed with the factors *age* (younger, older) and *feedback* (feedback, no feedback) as between-participants factors and *task type* (single, dual) as within-participants factor. All measures were averaged over set size.

Table B.1 presents the percentage of incorrect responses and misses in the dual-task condition as a function of segment and age group. Eighty trials were discarded for analysis due to an incomplete data set. As mentioned above, for the analysis of RT-VST, only VST trials in which a target was present and in which no error occurred were used. For the remaining trials ($N = 8820$), 24.6% was excluded due to errors. The mean percentages of errors in lane-change segments were 7.04% (incorrect responses older adults), 2.64% (incorrect responses younger adults), 4.81% (misses older adults) and 0.23% (misses younger adults), adding up to a total of 14.72%. In lane-keeping segments, the mean percentages of errors added up to a total of 9.87%, including incorrect responses older adults (6.93%), incorrect responses younger adults (1.73%), misses older adults (1.15%) and misses younger adults (0.06%) respectively.

For the VST single-task condition, from a total of 9000 trials, 1468 (16.30%) were discarded due to errors. All errors concerned incorrect responses. Older adults produced

Table B.1: Percentage of incorrect responses and misses in the dual-task condition as a function of segment and age group.

| Lane Change | | Lane Keeping | |
|---|---|---|---|
| Wrong button presses | | Wrong button presses | |
| Older Adults | 7.04 | Older Adults | 6.93 |
| Younger Adults | 2.64 | Younger Adults | 1.73 |
| Misses | | Misses | |
| Older Adults | 4.81 | Older Adults | 1.15 |
| Younger Adults | 0.23 | Younger Adults | 0.06 |

most incorrect responses (11.94%).

*RT-VST.* The ANOVA revealed a main effect of task type ($F$[1, 45] = 23.40, $\eta_p^2$ = .34, $p$ < .001): RT-VST in the single-task condition was significantly lower (M = 1153 ms) than that in the dual-task condition (M = 1254 ms). There was furthermore a main effect of age ($F$[1, 45] = 126.34, $\eta_p^2$ = .74, $p$ < .001) with RT-VSTs of older adults being significantly higher (M = 1386 ms) than that of younger participants (M = 1021 ms). No other significant main effects or interactions were found (all $p$s > .14).

*Incorrect Responses.* Figure B.10 presents the proportion of incorrect responses as a function of task type and age group. The ANOVA yielded a main effect of task type ($F$[1, 44] = 23.86, $\eta_p^2$ = .35, $p$ < .001): incorrect response rates were higher in the dual-task condition (M = .20) as compared to the single-task condition (M = .15). The ANOVA furthermore revealed a main effect of age ($F$[1, 44] = 72.72, $\eta_p^2$ = .62, $p$ < .001) with older adults producing more incorrect responses (M = .26) than younger adults (M = .08). The factors task type and age interacted significantly ($F$[1, 44] = 11.24, $\eta_p^2$ = .20, $p$ < .01): paired-samples $t$-tests showed that there was no statistical difference in incorrect responses for younger adults ($p$ > .11, one-tailed) whereas older adults suffered from an increase in task difficulty ($t$[23] = 5.33, $p$ < .001, one-tailed). The omnibus ANOVA showed no other significant main effects or interactions (all $p$s > .17).

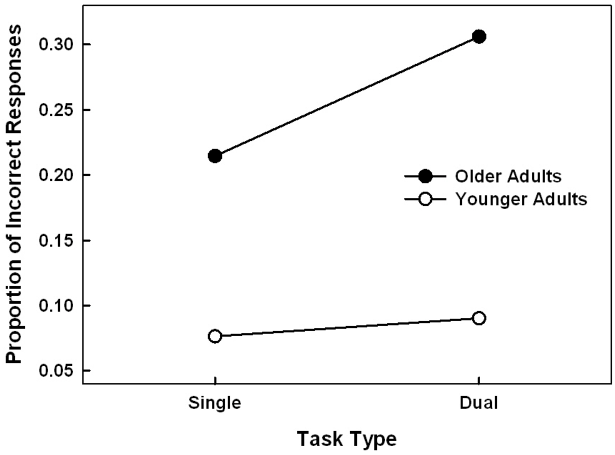


Figure B.10: Proportion of incorrect responses as a function of task type (single, dual) and age group (younger, older).

### B.1.3 Subjective Measures

A 3-way mixed-factors ANOVA taking into account the difference in single- and dual-task conditions (with the within-participants factor *task type*), age groups (*age*) and feedback condition (*feedback*) was conducted on the measure Mean RSME as well as each of the NASA-TLX scales.

*RSME.* Table B.2 presents means (M) and standard deviations (SD) on the RSME and each of the subscales of the NASA-TLX. The ANOVA only revealed a significant main effect of task type ($F[2, 92] = 56.81$, $\eta_p^2 = .55$, $p < .001$) on RSME: the dual-task condition provoked the highest ratings of mental demand, followed by the VST and the LCT.

*NASA-TLX.* Table B.2 presents the mean values of each of the subscales of the NASA-TLX and the effect of task type on each of these scales. The ANOVA revealed a main effect of the factor task type on all subscales except on the subscale performance. The LCT was in all cases rated as the least demanding, followed by the VST and then the dual-task condition.

Table B.2: Means (M), Standard Deviations (SD) and the effect of the factor task type for each of the three tasks (single-LCT, single-VST and dual-task) on the RSME and each of the subscales on the NASA-TLX.

| Scale | $M_{LCT}$(SD) | $M_{VST}$(SD) | $M_{DUAL}$(SD) | Effect Tasktype ($p$) |
|---|---|---|---|---|
| RSME | 29.15 (1.89) | 39.35 (2.34) | 53.36 ()3.40) | $<$ .001 |
| NASA-TLX | | | | |
| Mental Demand | 35.59 (3.37) | 54.42 (4.12) | 77.47 (4.90) | $<$ .001 |
| Physical Demand | 31.50 (3.36) | 30.03 (3.31) | 48.29 (4.65) | $<$ .001 |
| Temporary Demand | 39.22 (3.12) | 55.56 (4.59) | 69.97 (4.74) | $<$ .001 |
| Performance | 87.53 (3.44) | 86.90 (2.95) | 85.45 (3.24) | n.s. |
| Effort | 52.24 (4.08) | 68.86 (4.50) | 80.05 (4.97) | $<$ .001 |
| Frustration | 21.16 (3.00) | 33.20 (4.26) | 34.95 (4.37) | $<$ .001 |

Figure B.11 presents the significant interaction between task type and age for mean rating of Mental Demand on the NASA-TLX scale ($F[2, 92] = 3.74$, $\eta_p^2 = .08$, $p < .05$). A follow-up analysis with separate 1-way (task type) ANOVAs for each age group, revealed a simple main effect of task type for younger adults ($F[2, 46] = 55.83$, $\eta_p^2 = .71$, $p < .001$) as well as older adults ($F[2, 50] = 23.98$, $\eta_p^2 = .49$, $p < .001$). The interaction shown in Figure B.11 can thus be explained by the fact that ratings of mental demand increase for both age groups, with the increase being stronger for younger adults. This effect might be due to a general over-rating of the own performance frequently observed in older adults (see the Dunning-Kruger effect by Kruger & Dunning, 1999).

Figure B.12 presents Mean NASA-TLX ratings for the subscale Temporal Demand as a function of task type and feedback condition. As can be seen, the factor task type showed a tendency to interact with the factor feedback condition ($F[2, 92] = 3.02$, $\eta_p^2 = .06$, $p = .06$). The ratings of temporal demand seem to increase more or less parallel for the VST in both the FB- as well as the NOFB-condition as compared to the LCT condition. However in the dual-task condition, the subjective rating of temporal demand was much higher for the FB-condition as compared to the NOFB-condition. This might indicate that participants with feedback felt more stressed by the dual-task condition than participants without feedback.

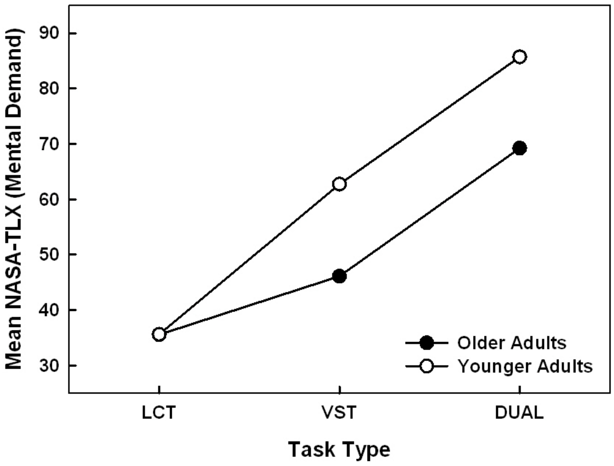


Figure B.11: Mean NASA-TLX ratings for Mental Demand as a function of task type (single-LCT, single-VST, dual-task) and age group (younger, older).

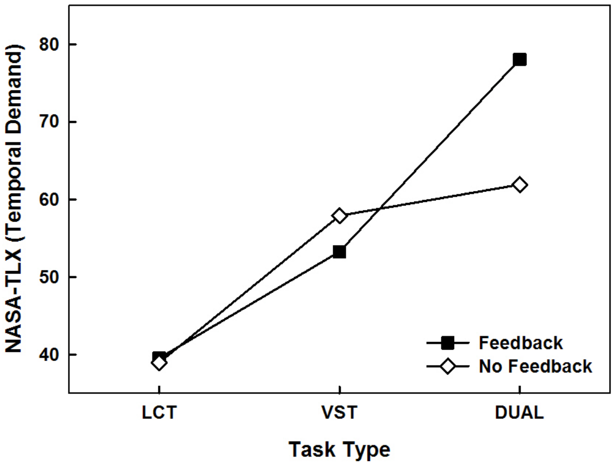


Figure B.12: Mean NASA-TLX ratings for Temporal Demand as a function of task type (single-LCT, single-VST, dual-task) and feedback condition (feedback, ne feedback).

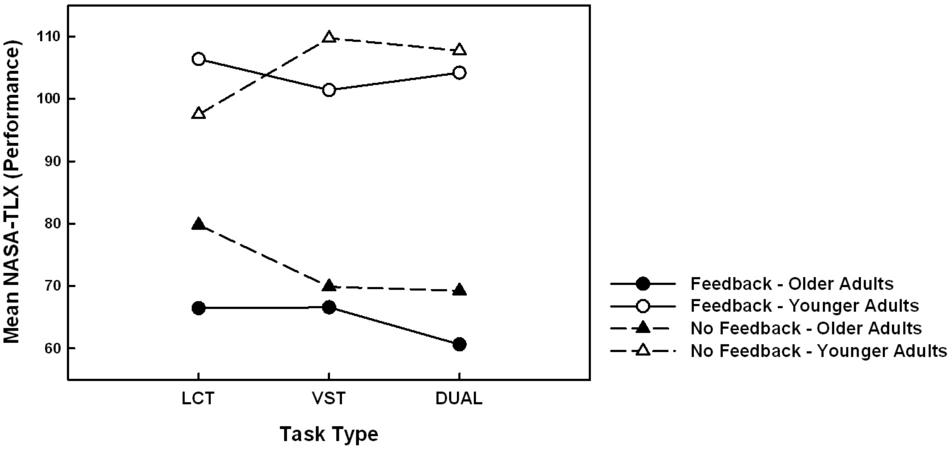


Figure B.13: Mean NASA-TLX ratings for Performance as a function of task type (single-LCT, single-VST, dual-task), feedback condition (feedback, no feedback) and age group (younger, older).

Figure B.13 presents Mean NASA-TLX ratings for the subscale Performance as a function of task type, feedback condition and age group. There was a tendency for those factors to interact ($F[2,\ 92] = 2.80$, $\eta_p^2 = .06$, $p = .07$) which indicates that not only younger adults rated their performance as worse than the older adults, but that this rating became more ”severe” in the absence of feedback. No other significant interactions could be found (all $p$s $> .11$).

## B.2 Analyzing Dual-Task Data in Detail

### B.2.1 Driving Measures

As in this section we have a closer look at dual-task data only, we are able to take into account the factor *block* (1-10), which takes into account potential learning effects within one session. Therefore, separate 3-way mixed-factor (*age* x *feedback* x *block*) ANOVAs were conducted on RT-LC, MT-LC and IRI. The driving measures MDEV, SDDEV, SDSW underwent separate 4-way ANOVAs taking into account the number of blocks: *age* x *feedback* x *segment* x *block*. To avoid partial redundancy with the data presented in the section above, only significant main effects and interactions including the factor block are reported here.

*RT-LC.* The ANOVA revealed a significant main effect of block ($F[9,\ 414] = 2.84$, $\eta_p^2 = .06$, $p < .01$) indicating that with practice within one session, all participants became faster at initiating a lane change ($\mathrm{M}_{\mathrm{Block1}} = 945$ ms; $\mathrm{M}_{\mathrm{Block10}} = 914$ ms). No other significant interactions with the factor block were found (all $p$s > .15).

*MT-LC.* The ANOVA revealed no significant main effect of block nor any interactions with the factor block for MT-LC (all $p$s > .47).

*IRI.* The ANOVA revealed a significant main effect of age on IRI ($F[1,\ 41] = 5.95$, $\eta_p^2 = .13$, $p < .05$): IRI of older adults was higher (M = 1423 ms) than that of younger adults (M = 1225 ms). This indicates that older adults suffered more from a secondary task than younger adults. The factor block just missed to reach significance ($F[9,\ 369] = 2.14$, $\eta_p^2 = .05$, $p = .05$) but indicates that with practice over 10 blocks, IRI tended to decrease ($\mathrm{M}_{\mathrm{Block1}} = 1381$ ms; $\mathrm{M}_{\mathrm{Block10}} = 1269$ ms). No other significant main effects or interactions could be found (all $p$s > .32).

*MDEV.* Figure B.14 presents MDEV as a function of block and feedback condition. There was no significant effect of block ($p = .11$), but there was a tendency for the factors block and feedback condition to interact ($F[9,\ 414] = 5.95$, $\eta_p^2 = .04$, $p = .08$). It seems that participants without feedback learned more over blocks than participants with feedback. The ANOVA revealed no other signficant interactions including the factor block (all $p$s > .12).

*SDDEV.* The 4-way interaction including the factors age, feedback, segment and block showed a tendency ($F[9,\ 414] = 1.92$, $\eta_p^2 = .04$, $p = .07$) but just missed to reach significance. No main effect nor any other interactions with the factor block were found (all $p$s > .17).

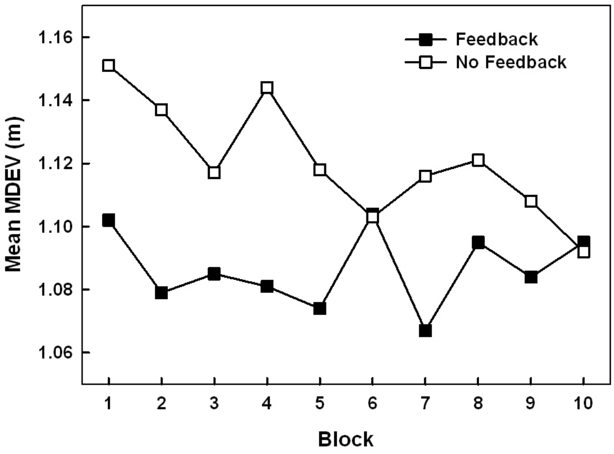


Figure B.14: Mean Lateral Deviation (MDEV) as a function of block (1-10) and feedback condition (feedback, no feedback).

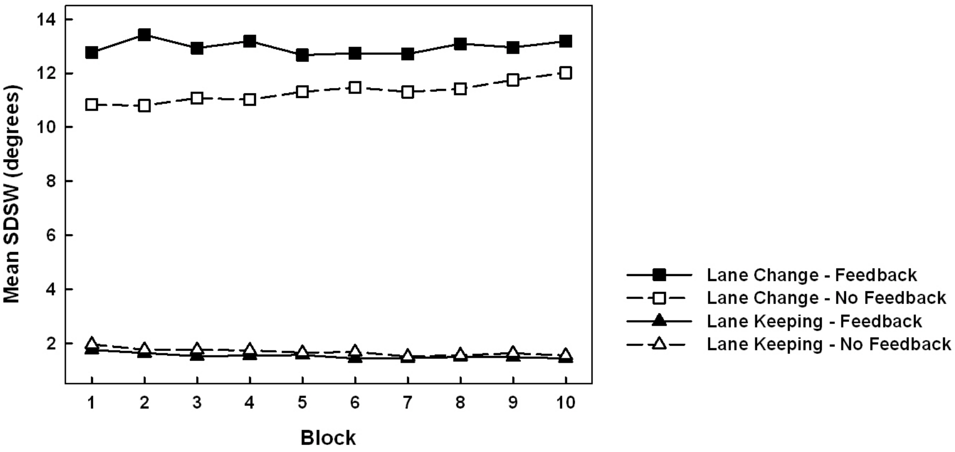


Figure B.15: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback).

*SDSW.* Figure B.15 presents mean SDSW as a function of block (1-10), segment and feedback condition. The 2-way interaction including the factors segment and block was significant ($F[9,\ 414] = 3.62$, $\eta_p^2 = .07$, $p < .01$). This 2-way interaction was however modulated by a significant 3-way interaction including the factors block, segment and feedback condition ($F[9,\ 414] = 2.35$, $\eta_p^2 = .05$, $p < .05$). To further analyze this interaction, we conducted two 2-way (block x segment) ANOVAs for the FB- and the NOFB-conditions respectively. The interaction between block and segment was significant for the NOFB-condition ($F[9,\ 216] = 4.72$, $\eta_p^2 = .16$, $p < .01$), but no significant interaction between those factors was found for the FB-condition ($p > .30$). The 3-way interaction can thus be explained by the fact that with practice, feedback had no effect on learning either in the LC and the LK segments. However, when no feedback was presented, the number of corrective steering-wheel movements increased for LC segments, but not for LK segments.

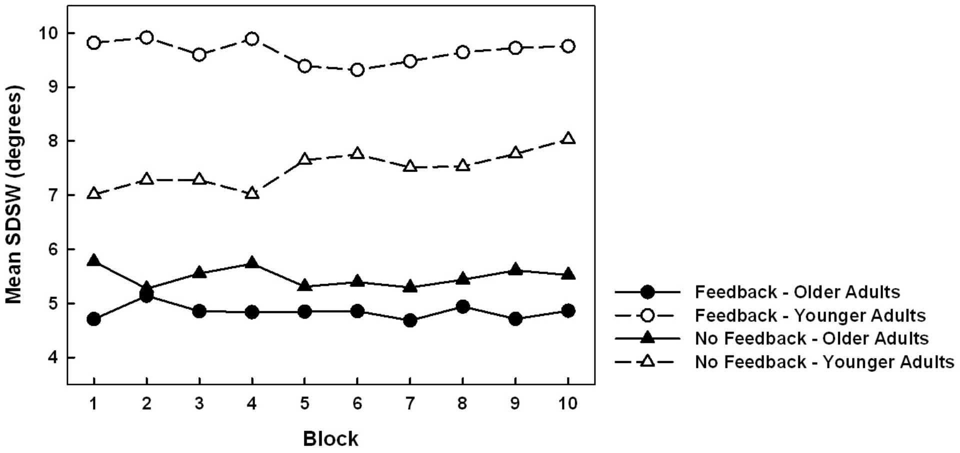


Figure B.16: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of block (1-10), feedback condition (feedback, no feedback) and age group (young, old).

Figure B.16 presents SDSW as a function of block, feedback condition and age group. The 3-way interaction including those factors showed a tendency ($F[9,\ 414] = 2.09$, $\eta_p^2 = .04$, $p = .08$) but missed to reach significance. It seems to indicate however, that practice had a beneficial effect on learning, with an increase of SDSW values over blocks, especially for younger adults when no feedback was provided. No other significant interactions with the factor block were observed (all $p$s $>$ .13).

### B.2.2 Visual Search Measures

In order to establish the influence of potential learning effects due to repetitions (*block*: 1-10) as well as the effect of segment (*segment*: LC, LK) as an additional factor to age group and feedback condition, separate 4-way mixed-factors (age x feedback x segment x block) ANOVAs were conducted on the following VST measures: RT-VST, incorrect responses and misses. All measures were averaged over set size. In order to avoid partial redundancy with results presented above, only relevant new interactions including the factors block and segment are reported here.

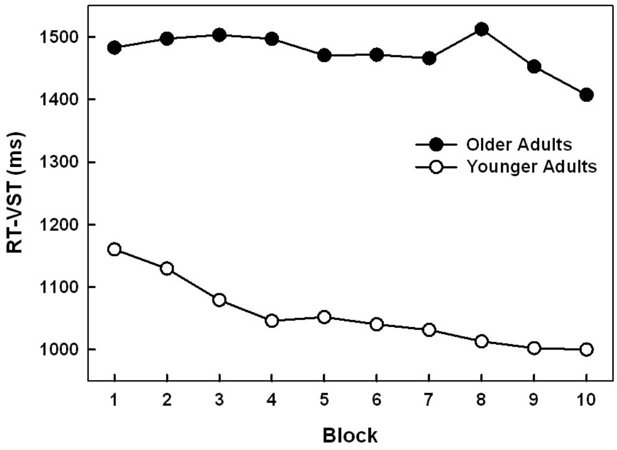
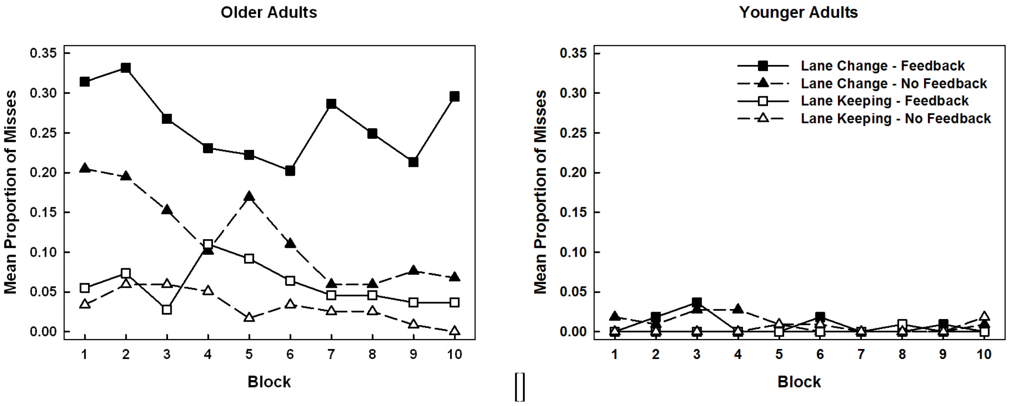


Figure B.17: Mean Reaction Time on the Visual Search Task (RT-VST) as a function of block (1-10) and age group (young, old).

*RT-VST.* The ANOVA yielded a significant main effect of segment ($F$[1, 44] $=$ 10.09, $\eta_p^2$ $=$ .20, $p$ $<$ .01): RT-VST was higher when performing a lane change (M $=$ 1311 ms) than when lane keeping (M $=$ 1220 ms). Figure B.17 presents mean RT-VST as a function of block and age group. There was a significant main effect of block ($F$[9, 369] = 4.64, $\eta_p^2 = .10$, $p < .001$): practice had a positive effect on RT-VST which decreased over blocks ($\mathrm{M}_{\mathrm{Block1}} = 1322$ ms; $\mathrm{M}_{\mathrm{Block10}} = 1204$ ms). The 2-way interaction including the factors block and age ($F$[9, 369] = 2.06, $\eta_p^2 = .05$, $p = .06$) just missed to reach significance. It indicates however that although there is a tendency for both age groups to benefit from practice over blocks, the benefits seem to be stronger for younger adults, resulting in lower VSTs.

*Incorrect Responses.* The ANOVA yielded a main effect of segment ($F$[1, 44] = 9.85, $\eta_p^2 = .18$, $p < .01$): the biggest proportion of incorrect responses occurred in the LC segment ($\mathrm{M}_{\mathrm{LC}} = .22$; $\mathrm{M}_{\mathrm{LK}} = .17$). No other main effects or interactions with the factors segment and block attained significance (all $p$s $> .10$).

*Misses.* Figure B.18a presents the proportion of misses as a function of block, segment and feedback condition for older adults. Figure B.18b presents the proportion of misses as a function of block, segment and feedback condition for younger adults. One can observe a main effect of block ($F$[9, 405] = 3.30, $\eta_p^2 = .07$, $p < .01$): practice had a beneficial effect on the number of misses ($\mathrm{M}_{\mathrm{Block1}} = .08$; $\mathrm{M}_{\mathrm{Block10}} = .05$). There was a main effect of segment as well ($F$[1, 45] = 21.77, $\eta_p^2 = .33$, $p < .001$) with most misses occurring in the LC segment ($\mathrm{M}_{\mathrm{LC}} = .19$, $\mathrm{M}_{\mathrm{LK}} = .06$). The ANOVA furthermore yielded a significant 2-way interaction between the factors block and age ($F$[9, 405] = 2.35, $\eta_p^2 < .05$, $p < .05$) as well as the factors block and segment ($F$[9, 405] = 2.15, $\eta_p^2 = .05$, $p < .05$) and the factors segment and age ($F$[1, 45] = 16.26, $\eta_p^2 = .27$, $p < .001$). The 3-way interaction including the factors block, segment and feedback condition just missed to reach significance ($F$[9, 405] = 2.10, $\eta_p^2 = .05$, $p = .05$). The 4-way interaction including the factors block, segment, age and feedback condition turned out to be significant however ($F$[9, 405] = 2.13, $\eta_p^2 = .05$, $p < .05$). To analyze this interaction, two separate 3-way ANOVAs including the factors block, segment and feedback condition were conducted for each age group. For the younger adults the 3-way interaction block x segment x feedback condition was not significant ($F$[9, 198] = 2.13, $\eta_p^2 = .05$, $p = .35$). For the older adults, this same interaction was significant ($F$[9, 207] = 2.38, $\eta_p^2 = .09$, $p < .05$). As can be seen in figure B.18a not only did older adults produce much more misses than younger adults, but feedback condition influenced the number of misses produced



[] []

Figure B.18: Panel (a): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for older adults. Panel (b): Proportion of misses as a function of block (1-10), segment (lane change, lane keeping) and feedback condition (feedback, no feedback) for younger adults.

in LC segments. When feedback was provided, the number of misses was higher, when no feedback was provided the number of misses was lower. This might indicate that older adults have difficulties focusing on both the LCT and the VST at the same time and when feedback forces them to prioritise the LCT, their VST suffers more. No other significant interactions including the factors block or segment were observed (all $ps >$ .24).

## B.2.3 Subjective Measures

A 3-way mixed-factors ANOVA taking into account the factor *block* (1-10) as well as the factors *feedback* (feedback, no feedback) and *age* (young, old) was conducted on the measure Mean RSME.

*Mean RSME.* The ANOVA did not reveal any significant main effect or interactions with the factor block (all $ps >$ .10).

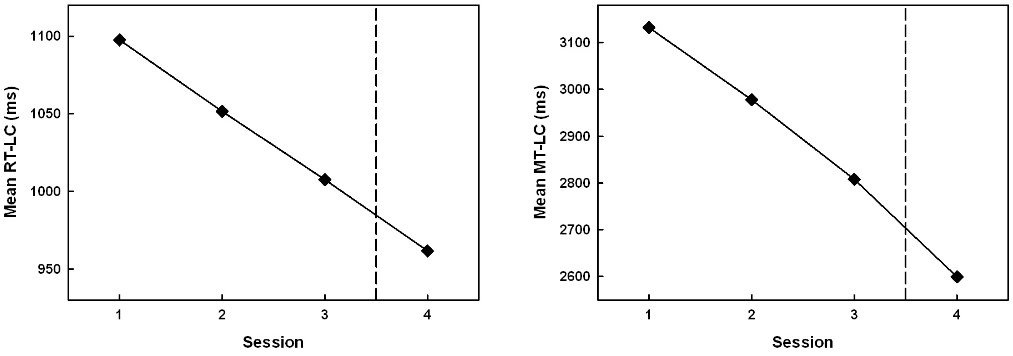
# C Results Experiment 3

Results are divided in two sections: first we analyzed the effect that adding a secondary task had on performance measures by comparing single- to dual-task data. Then we analyzed the dual-task condition more in detail. For both sections separately we examined driving, visual-search, dual-task, and subjective data separately. The different ANOVAs used for each section are described at the beginning of the section concerned. If a significant main effect or a significant simple main effect with the factor *session* occurred, $t$-tests were performed comparing the first and the third session to account for learning effects, and the third and the fourth session to account for retention effects. All $t$-tests were one-tailed. Note that due to the amount of data, figures were only included for significant or close to significant ($p =$.08) main effects or interactions. Note that close to significant interactions will only be described, without providing any exploratory follow-up tests.

## C.1 Comparing Single versus Dual-Task Data

### C.1.1 Driving Measures

Separate 3-way mixed-factors ANOVAs were conducted on the driving measures RT-LC and MT-LC. The ANOVA included the between-participants factor *age* (younger, older) to examine the effect of age of the participants, the within-participants factor *task type* (single-LCT, dual) to examine the effect of task complexity, and finally the effect of repetitions over sessions with the factor *session*. As described in the general method section (**??**), RT-LC and MT-LC measures were based on LC segments only. The driving measures MDEV, SDDEV, SDSW underwent separate 4-way ANOVAs including the between-participants factor *age* (younger, older) and the within-participants factors *task type* (single-LCT, dual), *session* (1-4) and finally *segment* (LC, LK) to take into



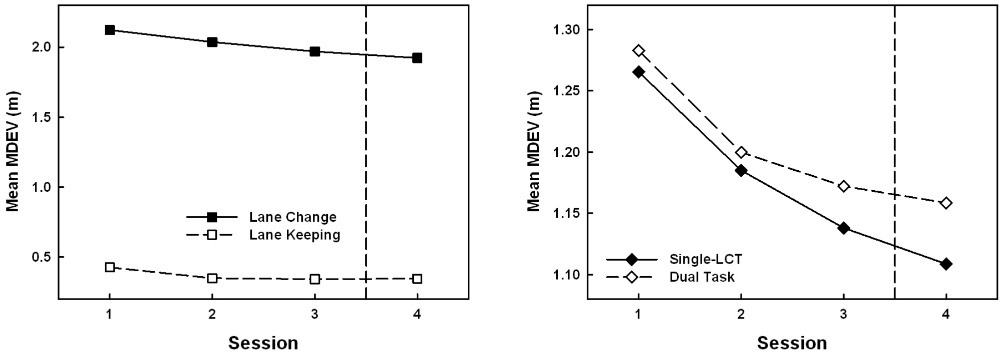
(a) (b)

Figure C.1: Reaction Time until Lane Change (a) and Movement Time (b) as a function of session (1-4).

account the effect of driving difficulty on those measures. All ANOVAs were averaged over the factor set size.

*RT-LC.* Figure C.2a presents RT-LC as a function of session. The ANOVA revealed a significant main effect of session on RT-LC ($F$[3, 54] = 9.52, $\eta_p^2 = .35$, $p < .001$) indicating that practice had a beneficial effect on reducing the onset time of the lane change. A follow-up analysis of this main effect with paired $t$-tests comparing the first and the third and the third and the fourth session, showed that learning took place between the first and the third session ($t$[19] = 2.56, $p < .05$), but that performance just missed to reach significance between the third and fourth session ($t$[19] = 1.91, $p$ = .07). The retention interval thus had statistically no effect on learning, but participants did not loose any acquired skills either. There was furthermore a significant main effect of task type ($F$[1, 18] = 8.74, $\eta_p^2 = .33$, $p < .01$): RT-LC was higher in dual-task conditions (M = 1048 ms) as compared to single-LCT conditions (M = 1011 ms). This indicates that adding a secondary task increases the reaction time until initiation of a lane change. No other significant main effects or interactions were observed (all $p$s > .39).

*MT-LC.* Figure C.2b presents MT-LC as a function of session. The ANOVA revealed a significant main effect of session on MT-LC ($F$[3, 54] = 9.01, $\eta_p^2 = .33$, $p < .01$): prac-



(a) (b)

Figure C.2: Mean Lateral Deviation (MDEV) as a function of (a) session (1-4) and segment (lane change, lane keeping) and (b) as a function of session (1-4) and task type (single-LCT, dual-task).

tice had a beneficial effect on reducing lane-change duration. Paired $t$-tests showed that there was an effect of learning between the first and the third session on MT-LC ($t[19] = 2.20$, $p < .05$), but, despite a tendency, the difference in performance between the third and fourth session was not significant ($t[19] = 2.06$, $p = .05$). Statistically speaking, the retention period thus did not improve learning, but it did not degrade previously learned skills either. There was furthermore a significant main effect of task type on MT-LC ($F[1,\ 18] = 25.50$, $\eta_p^2 = .59$, $p < .001$) with MT-LC being higher in dual-task conditions (M = 3028 ms) as compared to single-LCT conditions (M = 2730 ms). This indicates that the introduction of a secondary task had an effect on the movement time: lane-change periods become significantly longer. No other significant main effects or interactions were observed (all $p$s > .16).

*MDEV.* Figure C.2a presents mean MDEV as a function of session and segment. The ANOVA revealed a significant main effect of session ($F[3,\ 54] = 23.82$, $\eta_p^2 = .57$, $p < .001$) indicating that practice had a beneficial effect on lane-change accuracy as witnessed by a reduction in MDEV values. The ANOVA furthermore revealed a significant main effect of segment on MDEV ($F[1,\ 18] = 1179.43$, $\eta_p^2 = .99$, $p < .001$): MDEV in lane-change segments was higher (M = 2.01 m) than that of lane-keeping segments (M

= 0.37 m). The ANOVA revealed finally a significant main effect of age ($F$[1, 18] = 8.46, $\eta_p^2$ = .32, $p$ < .01) with MDEV values for older adults (M = 1.27 m) being higher than those of younger adults (M = 1.11 m). The factor task type showed a tendency but just missed to reach significance ($p$ > .08). The 2-way interaction including the factors session and segment was significant ($F$[3, 54] = 3.72, $\eta_p^2$ = .17, $p$ < .05). Follow-up tests with separate 1-way (session) ANOVAs for LC and LK segments respectively, revealed simple main effects for session in LC segments ($F$[3, 57] = 12.16, $\eta_p^2$ = .39, $p$ < .001) as well as LK segments ($F$[3, 57] = 9.94, $\eta_p^2$ = .34, $p$ < .01). The interaction comes from the fact that practice had a stronger effect on MDEV in LC as compared to LK segments. To analyze learning effects, follow-ups on the LC segments with paired-samples $t$-tests were performed, which showed that learning took place between the first and the third session ($t$[19] = 3.54, $p$ < .01), but that no difference in performance could be observed between the third and fourth session ($p$ > .20). The performance in the LK segments showed the same pattern: learning took place between the first and the third session ($t$[19] = 3.45, $p$ < .01), but no difference in performance could be observed between the third and fourth session ($p$ > .76). This means that in both segments, no learning took place within the retention interval, but performance did not decrease either.

Figure C.2b presents MDEV as a function of session and task type. The 2-way interaction between those two factors turned out to be significant as well ($F$[3, 54] = 3.27, $\eta_p^2$ = .15, $p$ < .05). One-way (session) ANOVAs for single-LCT and dual-task respectively revealed a significant simple main effect of session in both the single-LCT condition ($F$[3, 57] = 22.48, $\eta_p^2$ = .54, $p$ < .001) as well as the dual-task condition ($F$[3, 57] = 19.59, $\eta_p^2$ = .51, $p$ < .001), the interaction comes from the fact that with practice MDEV decreased stronger in single-LCT conditions as compared to dual-task conditions. An in-depth look at learning effects by analyzing the factor session with paired-samples $t$-tests, revealed that in the single-LCT condition, learning took place between the first and the third session ($t$[19] = 5.44, $p$ < .001), but that the retention interval had no effect on performance measures ($p$ > *.17*). The same pattern was observed for the dual-task condition, in which learning took place between the first and third session ($t$[19] = 5.38, $p$ < .001), but where the retention interval showed no effect on performance measures either ($p$ > .41). The ANOVA revealed a tendency for task type to interact with segment, but the interaction just missed to reach significance ($F$[1, 18] = 4.32, $\eta_p^2$

= .19, $p$ = .05). It seems to indicate however that task difficulty had a stronger effect on MDEV in LC as compared to LK segments.

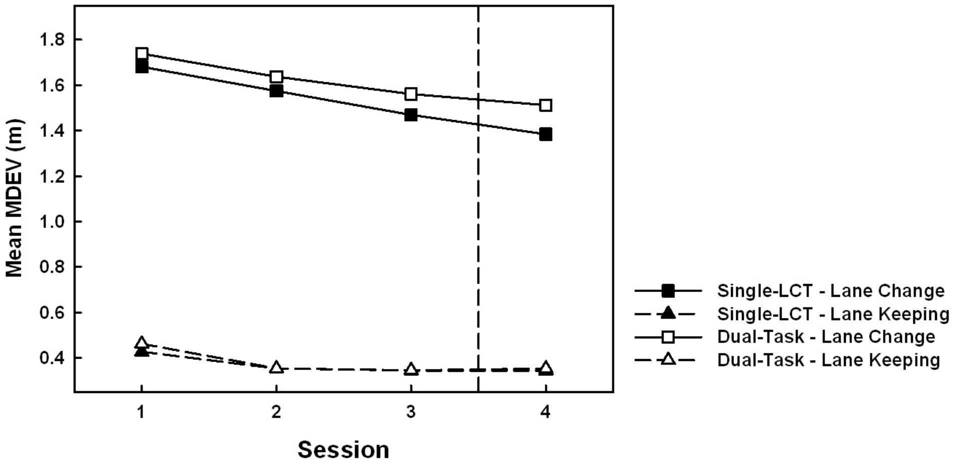


Figure C.3: Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping).

The 3-way interaction including session, task type and segment, graphically represented in Figure C.3, showed a tendency but missed to reach significance as well ($F$[3, 54] = 3.08 $\eta_p^2$ = .15, $p$ = .06). It might provide an indication though that practice had the strongest beneficial effect in LC segments and that the effect was strongest in the single-LCT condition.

Figure C.4 presents mean MDEV as a function of session, segment and age group. The 3-way ANOVA including those factors showed a tendency, but missed to reach significance ($F$[3, 54] = 3.01 $\eta_p^2$ = .14, $p$ = .06). It seems to indicate however that over sessions MDEV performance for all road segments and all age groups decreased. However, younger adults' MDEV decreased more in the LC segments, whereas the older adults benefited more from practice in the lane-keeping segments. No other significant interactions were observed (all $p$s $>$ .26).

*SDDEV.* Figure C.5 presents SDDEV as a function of session, segment and age. Al-

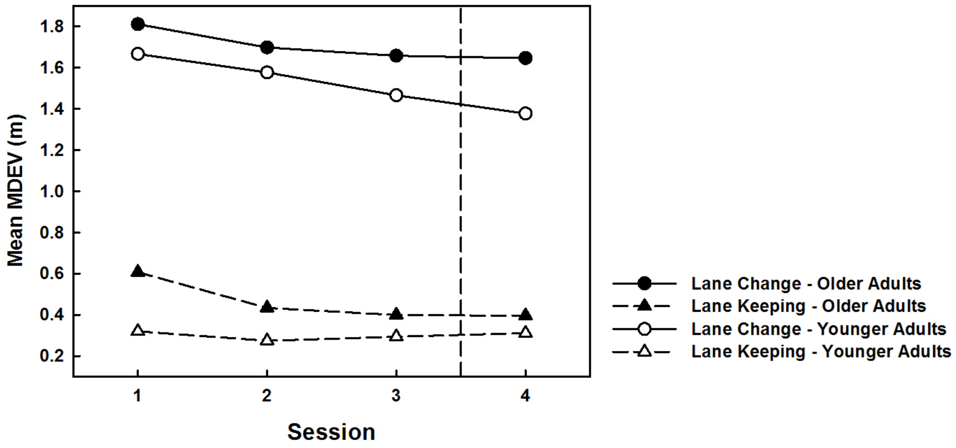


Figure C.4: Mean Deviation of Lateral Deviation (MDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older).

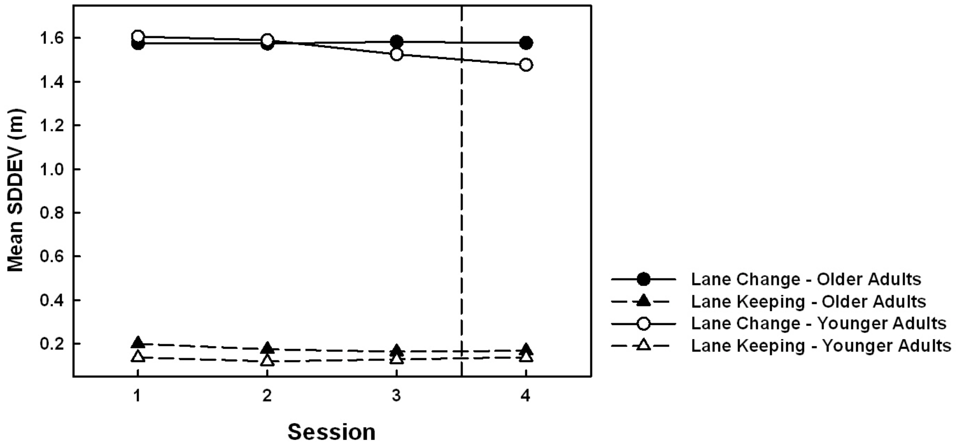


Figure C.5: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4), segment (lane change, lane keeping) and age (younger, older).

though difficult to observe in the figure, there was a significant main effect of session ($F[3, 54] = 7.60$, $\eta_p^2 = .30$, $p < .01$) on SDDEV with variability along the referece track decreasing with practice. There was furthermore a main effect of segment ($F[1, 18] = 2961.78$, $\eta_p^2 = .99$, $p < .001$): SDDEV values were higher in lane-change segments (M = 1.56 m) as compared to lane-keeping segments (M = 0.15 m). The interaction between session and age was significant ($F[1, 18] = 3.41$, $\eta_p^2 = .16$, $p < .05$), but modulated by a significant 3-way interaction including the factors session, segment and age ($F[3, 54] = 4.34$, $\eta_p^2 = .19$, $p < .05$). Follow-up post-hoc tests with two separate 2-way (session x age) ANOVAs for LC and LK segments respectively, revealed that there was no significant interaction between those factors for LK segments ($p > .11$), whereas for LC segments this interaction was significant ($F[3, 54] = 4.23$, $\eta_p^2 = .19$, $p < .05$). Separate one-way (session) ANOVAs for older and younger adults respectively, revealed that there was a simple main effect of session for younger adults ($F[3, 27] = 7.38$, $\eta_p^2 = .45$, $p < .01$), but not for older adults ($p > .34$). Only younger adults' performance thus improved with practice in the LC segments. An in-depth analysis of this improvement with paired-samples $t$-tests revealed that indeed learning took place between the first and the third session ($t[9] = 2.71$, $p < .05$), but that SDDEV values were the same between the third and fourth session ($p > .25$). In other words, the retention interval had no effect on learning, neither did it affect any previously acquired skills.

Figure C.6 presents SDDEV as a function of task type and segment. The ANOVA revealed a significant interaction between those two factors ($F[1, 18] = 7.35$, $\eta_p^2 = .29$, $p < .05$). Follow-up tests with paired-samples $t$-tests showed that SDDEV values differed significantly between single-LCT- and dual-task conditions for LC-segments ($t[19] = 1.95$, $p < .05$) as well as LK segments ($t[19] = 3.01$, $p < .01$). Looking at Figure C.6, the interaction can be explained by the fact that the factor task type has different effects according to the driving segment analyzed: with an increase in task difficulty, SDDEV values in the LC segments increase whereas SDDEV values in LK segments decrease.

Figure C.7 presents mean SDDEV as a function of session and task type. The 2-way interaction between those factors was significant ($F[3, 54] = 5.82$, $\eta_p^2 = .24$, $p < .01$). A follow-up of this interaction with two separate 1-way (session) ANOVAs for each task type respectively, revealed a significant simple main effect of session in single-LCT con-

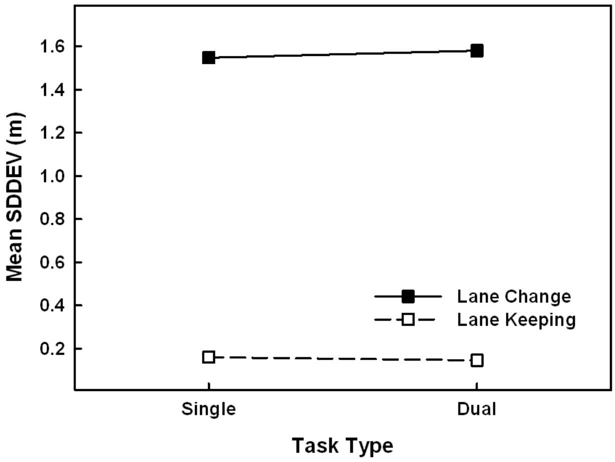


Figure C.6: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of task type (single-LCT, dual) and segment (lane change, lane keeping).

ditions ($F[3, 57] = 8.67$, $\eta_p^2 = .31$, $p < .001$), as well as dual-task conditions ($F[3, 57] = 3.46$, $\eta_p^2 = .15$, $p < .05$). However, as can be seen in the figure, practice had a greater beneficial effect on single-LCT conditions as compared to dual-task conditions. An in-depth analysis of the effect of session for both conditions with paired-samples $t$-tests, revealed that learning took place between the first and third session in the single-LCT condition ($t[19] = 3.12$, $p < .01$, ), but not in the dual-task condition ($t[19] = 1.97$, $p = .06$). The retention interval (analyzed by comparing the third and fourth session with paired-samples $t$-tests) had no effect on either conditions (all $p$s $> .18$).

The 3-way interaction including the factors session, task type and segment showed a tendency, but just missed to reach significance ($F[3, 54] = 2.54$, $\eta_p^2 = .12$, $p = .09$). It seems to indicate however that practice had no beneficial effect in LK segments (both for single-LCT task settings as well as dual-task settings), but that in LC segments, performance improved more in single-LCT conditions as compared to dual-task conditions. The omnibus ANOVA revealed no other significant interactions (all $p$s $> .15$).

*SDSW.* Figure C.8 presents SDSW as a function of session, task type and segment. The ANOVA revealed a main effect of task type ($F[1, 18] = 34.81$, $\eta_p^2 = .66$, $p < .001$):

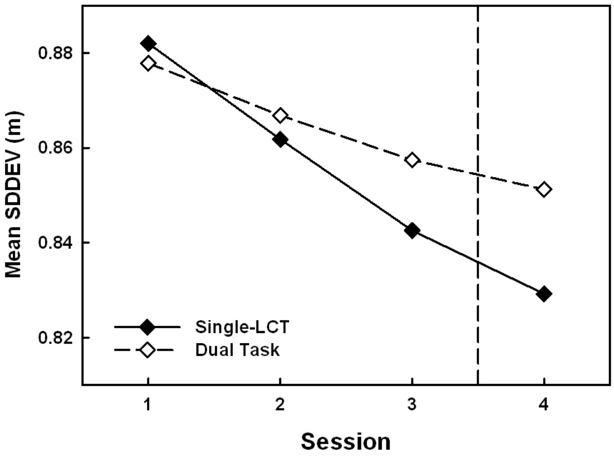


Figure C.7: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4) and task type (single-LCT, dual).

SDSW was higher in single-LCT conditions (M = 5.13°) as compared to dual-task conditions (M = 3.93°). The ANOVA furthermore revealed a main effect of segment ($F$[1, 18] = 75.22, $\eta_p^2$ = .81, $p < .001$) with SDSW values being significantly higher in LC (M = 7.91°) as compared to LK segments (1.15°). There was a main effect of session as well ($F$[3, 54] = 8.03, $\eta_p^2$ = .31, $p < .01$): with practice, SDSW values increased. Although not directly visible in figure C.8, the 2-way interaction between session and segment was significant ($F$[3, 54] = 10.48, $\eta_p^2$ = .37, $p < .001$) just like the 2-way interaction between session and task type ($F$[3, 54] = 4.42, $\eta_p^2$ = .20, $p < .05$) as well as the 2-way interaction between task type and segment ($F$[1, 18] = 22.31, $\eta_p^2$ = .55, $p < .001$). All these 2-way interactions were however modulated by a significant 3-way interaction including the factors session, task type and segment ($F$[3, 54] = 4.89, $\eta_p^2$ = .21, $p <$ .05). To further follow up on this interaction we conducted two separate 2-way (session x task type) ANOVAs for LC and LK segments respectively. There was no significant interaction between the factors session and task type in LK segments ($p > .39$), but the interaction reached significance in LC segments ($F$[3, 57] = 4.79, $\eta_p^2$ = .20, $p < .05$). Separate 1-way (session) ANOVAs for single-LCT and dual-task conditions respectively showed that there was both a significant simple main effect of session in the single-LCT condition ($F$[3, 57] = 7.67, $\eta_p^2$ = .29, $p < .01$) as well as the dual-task condition ($F$[3, 57]

$= 6.45$, $\eta_p^2 = .25$, $p < .01$). As can be seen in Figure C.8, practice had a stronger effect on single-LCT conditions as compared to dual-task conditions. An in-depth analysis of the session effects with paired-samples *t*-tests however, showed that some learning took place between the first and the third session in the dual-task condition ($t[19] = 2.26$, $p < .05$), but in the single-LCT condition only showed a tendency ($t[19] = 2.04$, $p = .06$). The retention interval only showed a tendency in the single-LCT condition ($p = .09$), but no significant effect in the dual-task condition ($p > .32$).

Finally, the 2-way interaction between segment and age showed a tendency, but missed to reach significance ($F[1, 18] = 3.19$, $\eta_p^2 = .15$, $p = .09$). The omnibus ANOVA revealed no other main effects or interactions (all *p*s $> .12$).

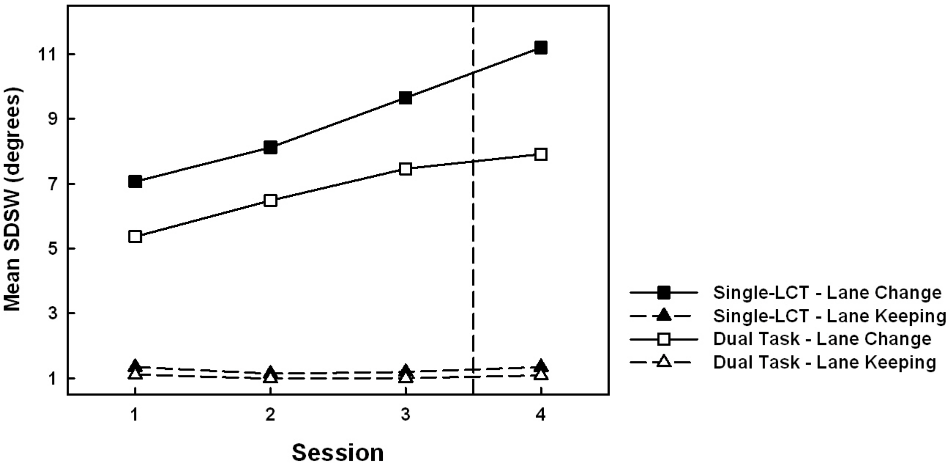


Figure C.8: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of session (1-4), task type (single-LCT, dual) and segment (lane change, lane keeping).

## C.1.2 Visual Search Measures

Performance on the VST was analyzed by the RT-VST as well as computing the proportion of incorrect responses for each condition and participant. As no misses occurred in the single-VST condition, no ANOVAs were performed for the measure misses. Values for incorrect responses were arcsine transformed to deal with the non-normality of proportions (e.g., see Winer, 1971). Separate 4-way mixed-factors ANOVAs were conducted on the visual-search measures RT-VST and the number of incorrect responses. All ANOVAs included the between-participants factor *age* (younger, older) to establish the effect of age, the within-participants factor *task type* (single-VST, dual) to examine the effect of task complexity, the effect of practice with the factor *session* and finally the effect of secondary-task complexity with the factor *set size*.

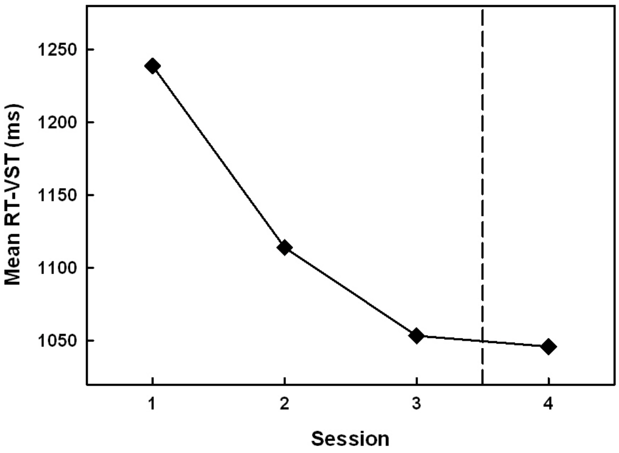


Figure C.9: Reaction Time on the Visual Search Task (RT-VST) as a function of session (1-4).

*RT-VST.* The ANOVA revealed a significant main effect of task type on RT-VST ($F[1, 17] = 35.56$, $\eta_p^2 = .68$, $p < .001$): RT-VSTs were higher in the dual-task condition (M = 1167 ms) as compared to the single-VST condition (M = 1060 ms). The ANOVA revealed a main effect of session as well ($F[3, 51] = 59.75$, $\eta_p^2 = .78$, $p < .001$). Figure C.9 presents RT-VST as a function of session and it can be observed that practice had a beneficial effect on reducing RT measures on the VST. A follow-up on the learning and retention effects with paired-samples *t*-tests showed that learning took place between the first and the third session ($t[19] = 9.97$, $p < .001$), but that the retention period had no effect (RT-VST values between the third and fourth session did not differ) ($p >$

.59). The main effect of set size was significant as well ($F$[2, 34] = 265.86, $\eta_p^2$ = .94, $p$ < .001): the search display with 25 items yielded the highest reaction times (M = 1218 ms), followed by the display with 16 items (M = 1109 ms) and the display with 9 items (M = 1012 ms). And finally there was a main effect of age on RT-VST ($F$[1, 17] = 55.90, $\eta_p^2$ = .77, $p$ < .001). Older adults needed more time to respond to a visual-search task (M = 1302 ms) as compared to younger adults (M = 924 ms). The interaction between set size and age showed a tendency, but missed to reach significance ($F$[2, 34] = 2.93, $\eta_p^2$ = .15, $p$ = .09). No other significant interactions were found (all $p$s > .10).

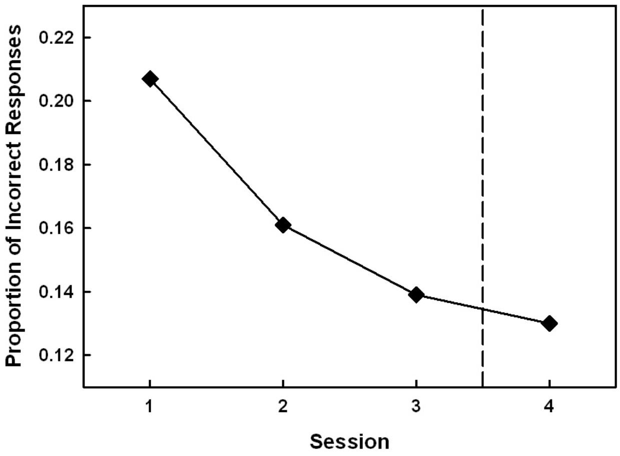


Figure C.10: Proportion of incorrect responses as a function of session (1-4).

*Incorrect Responses.* The ANOVA revealed a significant main effect of task type on the proportion of incorrect responses ($F$[1, 17] = 29.21, $\eta_p^2$ = .63, $p$ < .001): the proportion of incorrect responses was higher in the dual-task condition (M = .40) as compared to the single-VST condition (M = .34). The ANOVA revealed a main effect of session as well ($F$[3, 51] = 29.55, $\eta_p^2$ = .64, $p$ < .001). Figure C.10 presents the proportion of incorrect responses as a function of session in which it can be observed that practice had a beneficial effect on the proportion of incorrect responses produced in the visual search task. A follow-up analysis of this interaction showed that the proportion of incorrect responses significantly decreased between the first and the third session, indicating a learning effect ($t$[19] = 7.93, $p$ < .001). However, although a numerical difference between the third (M = .14) and fourth (M = .13) session could be observed, this difference was statistically not significant ($p > .40$). The retention period had thus no effect on the proportion of errors produced. The main effect of set size was significant as well ($F$[2, 34] = 86.23, $\eta_p^2 = .84$, $p < .001$) with the 25-items search display yielding the highest proportion of incorrect responses (M = .23), followed by the 16-items display (M = .15) and finally the 9-items display (M = .10). There was finally a main effect of age on the proportion of incorrect responses ($F$[1, 17] = 47.73, $\eta_p^2 = .74$, $p < .001$) with older adults producing a higher proportion of incorrect responses (M = .25) as compared to younger adults (M = .07).

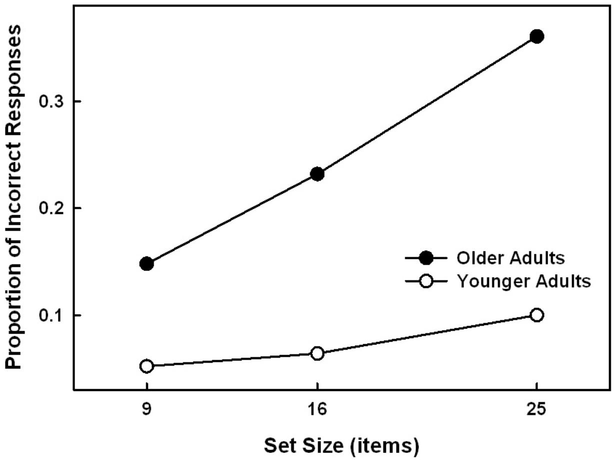


Figure C.11: Proportion of incorrect responses as a function of set size (9, 16, 25) and age group (young, old).

Figure C.11 presents the proportion of incorrect responses as a function of set size and age group. The observed the 2-way interaction between those factors was significant ($F$[2, 34] = 17.83, $\eta_p^2 = .51$, $p < .001$). Follow-up analysis with two separate 1-way (set size) ANOVAs for younger and older adults respectively, revealed that there was a simple main effect of set size for younger ($F$[2, 16] = 14.26, $\eta_p^2 = .64$, $p < .001$) as well as for older adults ($F$[2, 18] = 92.58, $\eta_p^2 = .91$, $p < .001$). As can be seen in Figure C.11, the increase in the proportion of incorrect responses was stronger with an increase in set size for older adults as compared to younger adults. This might indicate that older adults "suffer" more from an increase in set size than younger adults. The factor task type tended to interact with the factor age ($F$[1, 17] $= 7.42$, $\eta_p^2 = .30$, $p < .05$) but missed to reach significance. It might provide an indication however that the increase in incorrect responses when increasing task difficulty was stronger for older adults than for younger adults.

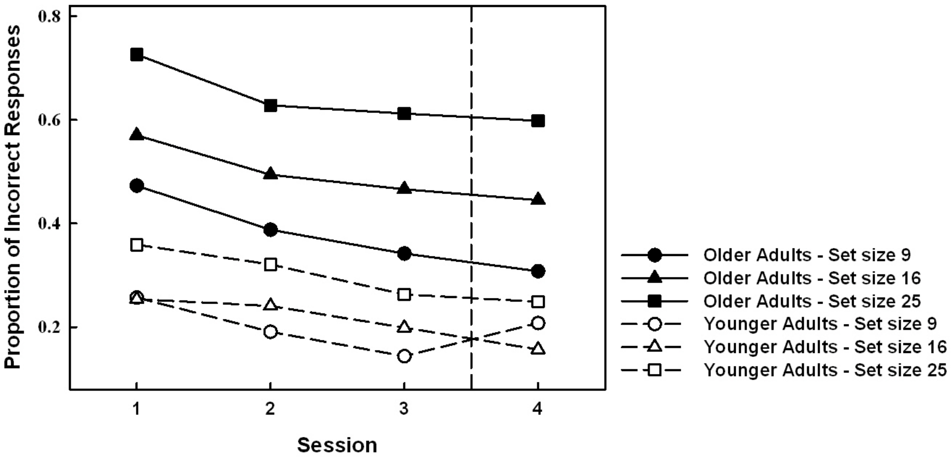
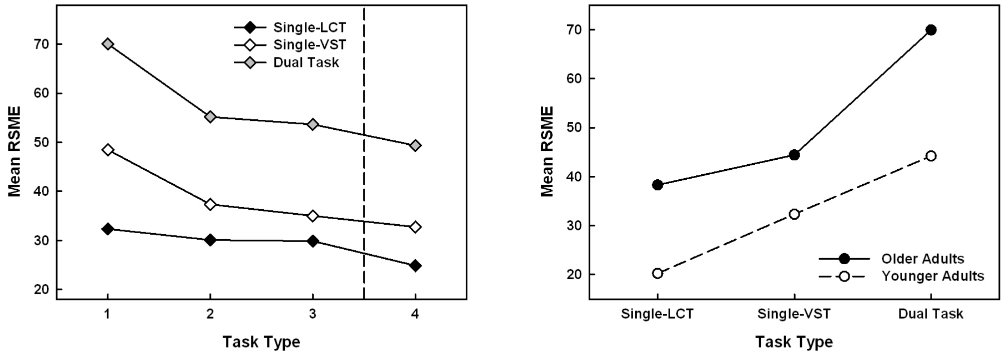


Figure C.12: Proportion of incorrect responses as a function of session (1-4), age group (young, old) and set size (9, 16, 25).

Figure C.12 presents the proportion of incorrect responses as a function of session, age group and set size. The 3-way interaction between those factors showed a tendency, but missed to reach significance ($F$[6, 102] $= 2.23$, $\eta_p^2 = .12$, $p = .07$) as well. Looking at the figure, it might however provide an indication that practice has a different effect on older adults than it has on younger adults. No other interactions reached significance (all $p$s $> .21$).

## C.2 Subjective Measures

Subjective measures assessed with the RSME and the NASA-TLX rating scales were analyzed with 3-way mixed-factors ANOVAs taking into account the between-participants



(a) (b)

Figure C.13: Mean Rating Scale of Mental Effort (RSME) as a function of (a) session (1-4) and task type (single-LCT, single-VST, DUAL) and (b) task type (single-LCT, single-VST, DUAL) and age group (younger, older).

factor *age* (younger, older) and the within-participants factor *task type* (single-LCT, single-VST, dual) to account for differences in mental demand rating according to task complexity as well as the within-participants factor *session* (1-4) to assess the influence of practice on mental demand ratings.

*RSME.* Figure C.13a shows mean RSME as a function of session and task type. The ANOVA yielded a main effect of task type on the subjective rating of mental demand ($F[2, 36] = 59.52$, $\eta_p^2 = .77$, $p < .001$): the dual-task condition was subjectively rated the most mentally demanding (M = 57.06), followed by the single-VST (M = 38.38) and the single-LCT (M = 29.26). The ANOVA showed furthermore a significant main effect of session ($F[3, 54] = 12.53$, $\eta_p^2 = .41$, $p < .001$): with practice over sessions the subjective rating of mental demand decreased for all tasks. Finally there was a significant main effect of age ($F[1, 18] = 8.63$, $\eta_p^2 = .32$, $p < .01$) with older adults rating all task conditions as mentally more demanding than younger adults. There was a significant 2-way interaction between session and task type ($F[6, 108] = 4.53$, $\eta_p^2 = .20$, p $< .01$): with practice, all tasks were rated less mentally demanding, but the decrease was strongest for the dual-task condition. Figure C.13b shows mean RSME as a function of task type and age group. One can observe a significant 2-way interaction between task type and age ($F[2,\ 36] = 3.45$, $\eta_p^2 = .16$, $p < .05$). Follow-up analysis with two separate 1-way (task type) ANOVAs for younger and older adults respectively, revealed that there was a main simple effect of task type for younger ($F[2,\ 18] = 24.32$, $\eta_p^2 = .73$, $p < .001$) as well as older adults ($F[2,\ 18] = 37.06$, $\eta_p^2 = .81$, $p < .001$). As can be seen in Figure C.13b, the increase in subjective mental effort increased stronger for older adults going from the single-LCT to the dual-task condition as compared to younger adults. This might indicate that older adults "suffered" more from a dual-task condition than zyounger adults.

Table C.1: Results of 3-way (session x task type x age) mixed-factors ANOVAs performed on each of the NASA-TLX subscale ratings. Presented are each of the NASA-TLX subscales, the factors that reached significance or close-to-significance on each of those scales, the degrees of freedom (df), $F$-values, effect sizes ($\eta_p^2$) and probability levels for statistical significance ($p$).

| NASA-TLX Subscale | Factor | df | $F$ | $\eta_p^2$ | $p$ |
|---|---|---|---|---|---|
| Mental Demand | Task Type | 2,36 | 21.74 | .55 | $< .001$* |
| | Session | 3,54 | 12.48 | .41 | $< .001$* |
| Physical Demand | Task Type | 2,36 | 14.65 | .45 | $< .001$* |
| | Age | 1,18 | 12.29 | .41 | $< .01$* |
| Temporal Demand | Task Type | 2,36 | 16.72 | .48 | $< .001$* |
| | Session | 3,54 | 4.24 | .19 | $< .05$* |
| | Age | 1,18 | 3.61 | .17 | $= .07$ |
| Performance | Task Type | 2,36 | 2.91 | .14 | $= .07$ |
| | Age | 1,18 | 7.16 | .28 | $< .05$* |
| Effort | Task Type | 2,36 | 29.66 | .62 | $< .001$* |
| | Session | 3,54 | 5.50 | .23 | $< .01$* |
| Frustration | Task Type | 2,36 | 7387 | .30 | $< .01$* |
| | Session | 3,54 | 3.94 | .18 | $< .05$* |

*NASA-TLX.* Table C.1 presents the results of the 3-way mixed-factors ANOVA on the 6 subscales of the NASA-TLX. One can observe that the factor task type had an effect on all subscales but performance. Participants subjectively rated the dual-task condition as the most mentally, physically and temporary demanding, as well as demanding most effort and generating most frustration. The dual-task condition was followed by the VST on those 5 subscales and the LCT was rated as the least mentally, physically and temporarily demanding, as well as demanding the least effort and provoking the least frustration. Although not significant, the subscale performance showed a tendency in the same direction. The factor session had a significant main effect on ratings of mental demand, temporal demand, effort and frustration: practice significantly reduced subjective ratings for those aspects of mental task demand. There were finally some age effects on the subscales of physical demand, temporal demand and performance. As to the ratings of physical and temporal demand, older adults rated both scales significantly higher than younger adults. As to the subscale performance, younger adults generated higher values, which for this particular subscale, means that they rated their own performance more severely than older adults did. All other main effects and interactions did not reach significance (all $p$s $> .17$).

## C.3 Analyzing Dual-Task Data in Detail

### C.3.1 Driving Measures

In this section we have a closer look at dual-task data. Separate 3-way mixed-factors ANOVAs were conducted on the driving measures RT-LC and MT-LC. The ANOVA included the within-participants factor *set size* (9, 16, 25) taking into account the effect of secondary task difficulty, the within-participants factor *session* (1-4) to analyze the effect of repetition over sessions and finally the between-participants factor *age* (younger, older) to account for age effects. As described in the general method section (**??**), RT-LC and MT-LC measures were based on LC segments only. The driving measures MDEV, SDDEV, SDSW underwent separate 4-way ANOVAs taking into account the number of sessions, the driving task difficulty, secondary task difficulty and age group: *session* x *segment* x *set size* x *age*. To avoid partial redundancy with the data presented in the section above, only relevant new main effects and interactions including the factor set size are reported here.

*RT-LC.* The ANOVA revealed no significant main effect nor any interactions with the factor set size (all $p$s $> .32$).

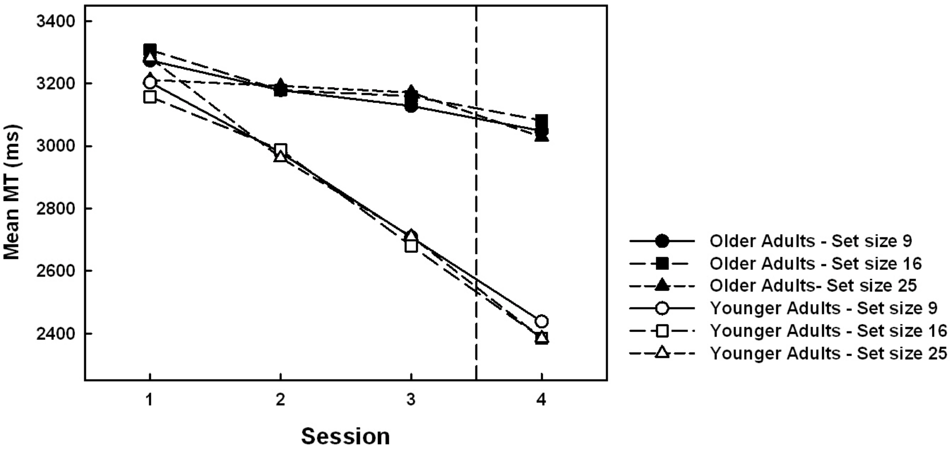


Figure C.14: Mean Movement Time (MT-LC) as a function of session (1-4), age group (younger, older) and set size (9, 16, 25).

*MT-LC*. Figure C.14 presents MT-LC as a function of session, age group and set size. The 2-way interaction between set size and age ($F[3,\ 51] = 3.20$, $\eta_p^2 = .16$, $p = .07$) showed a tendency, but just missed to reach significance. This close-to-significant interaction was however modulated by a significant 3-way interaction between the factors session, set size and age ($F[6,\ 102] = 2.64$, $\eta_p^2 = .13$, $p < .05$). To follow up on this interaction two separate 2-way (session x set size) ANOVAs were conducted for each age group respectively. There was no significant interaction between session and set size for older adults ($p > .38$) and this same interaction only showed a tendency for younger adults ($F[6,\ 48] = 2.36$, $\eta_p^2 = .23$, $p = .09$). The interaction can thus be explained by the fact that practice had no effect on movement times of older adults, independent of the set size presented, but that for younger adults movement times tended to decrease with practice. No significant main effect of set size or other interactions with this factor were found (all $p$s $> .31$).

*MDEV*. No significant main effect of the factor set size as well as significant interactions with this factor were found (all $p$s $> .12$). This indicates that secondary task difficulty had no significant effect on MDEV.

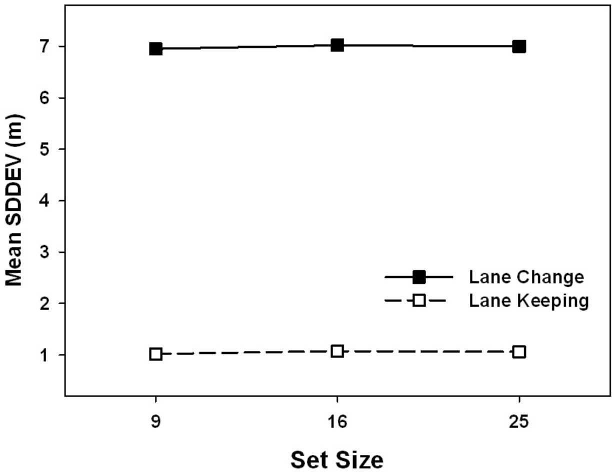


Figure C.15: Standard Deviation of Lateral Deviation (SDDEV) as a function of set size (9, 16, 25) and segment (lane change, lane keeping).

*SDDEV.* Figure C.15 represent SDDEV as a function of set size (9, 16, 25) and segment (LC, LK). The main effect of set size showed a tendency ($F[2,\ 34] = 2.69$, $\eta_p^2 = .14$, $p = .09$) but missed to reach significance. There was a significant 2-way interaction however between the factors set size and segment ($F[2,\ 34] = 3.94$, $\eta_p^2 = .19$, $p < .05$). To follow up on this interaction, two separate 1-way (set size) ANOVAs were conducted for LC and LK segments respectively. The ANOVA yielded a significant simple main effect of set size for LK segments ($F[2,\ 36] = 9.95$, $\eta_p^2 = .36$, $p < .01$) but this effect only showed a tendency for LC segments ($F[2,\ 36] = 3.21$, $\eta_p^2 = .15$, $p = .06$). This means that with increasing set size SDDEV values changed in LK segments, but not in LC segments. All other interactions including the factor set size did not reach significance (all $p$s $> .10$).

*SDSW.* No significant main effect of set size nor significant interactions with this factor were found (all $p$s $> .14$). This indicates that secondary task difficulty had no effect on SDSW.

## C.3.2 Visual-Search Measures

Separate 4-way mixed-factors ANOVAs were conducted on the visual-search measures RT-VST, the proportion of incorrect responses and the proportion of misses. The values concerning the proportion of incorrect responses and misses were arcsine transformed to deal with the non-normality of proportions (e.g., see Winer, 1971). All ANOVAs included the between-participants factor *age* (younger, older) to examine the effect of age of the participants, the effect of repetitions over sessions with the factor *session*, the effect of driving task difficulty with the factor *segment* (LC, LK) and finally the effect of secondary-task complexity with the factor *set size* (9, 16, 25).

To avoid redundancy, only relevant main effects or interactions with the factor segment are reported for RT-VST and wrong button presses. However, as the visual-search measure "misses" has not been analyzed this far, the main effects and interactions for all factors are reported in this section.

*RT-VST.* Figure C.16 presents the results for mean RT-VST as a function of session and segment. The 2-way interaction between session and segment was significant ($F$[3, 51] = 3.99, $\eta_p^2$ = .19, $p < .05$). A follow-up on this significant interaction with two separate 1-way (session) ANOVAs for each segment, revealed a significant simple main effect of session for both LC segments ($F$[3, 54] = 31.09, $\eta_p^2$ = .63, $p < .001$) as well as LK segments ($F$[3, 54] = 31.60, $\eta_p^2$ = .64, $p < .001$). The interaction comes from the observation that RT-VSTs in LC segments decreased more with practice than RT-VSTs in LK segments. A follow-up on those session effects with paired-samples $t$-tests showed that learning took place between the first and the third session for both LC ($t$[19] = 6.63, $p < .001$) as well as LK segments ($t$[19] = 6.58, $p < .001$). However, the retention interval had no effect on RT-VST: there was no significant difference between RT-VST in the third and fourth session for LC ($p > .19$) as well as LK segments ($p > .35$). No significant main effect of the factor segment nor any interactions with this factor were found (all $p$s $> .25$).

*Incorrect Responses.* The ANOVA revealed a main effect of segment on the proportion of incorrect responses ($F$[1, 17] = 21.40, $\eta_p^2$ = .56, $p < .001$): more incorrect responses occurred for visual search tasks that appeared while lane changing (M = .21) as com-

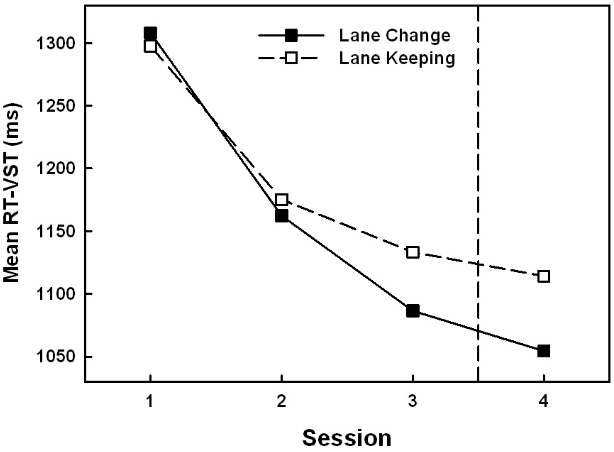


Figure C.16: Mean Reaction Time on the Visual Search Task (RT-VST) as a function of session (1-4) and segment (lane change, lane keeping).

pared to lane keeping (M = .15). The interaction between segment and age showed a tendency, but missed to reach significance ($F$[1, 17] = 3.96, $\eta_p^2$ = .19, $p$ = .06). It indicates however that the proportion of incorrect responses for older adults was higher when a visual search task was presented in LC segments as compared to presentation in LK segments. No other interactions with the factor segment reached significance (all $p$s > .13).

*Misses.* Figure C.17 presents the number of misses as a function of session, segment and age group. The ANOVA revealed a significant main effect of segment on the proportion of misses produced ($F$[1, 17] = 7.73, $\eta_p^2$ = .31, $p < .05$) with the proportion of misses being significantly higher in the lane-change segments (M = .04) as compared to the lane-keeping segments (M = .01). There was furthermore a significant main effect of session on the proportion of misses ($F$[3, 51] = 19.78, $\eta_p^2$ = .54, $p < .001$): repetition over sessions reduced the proportion of misses produced. Finally there was a main effect of age ($F$[1, 17] = 11.17, $\eta_p^2$ = .40, $p < .01$). The proportion of misses produced by older adults was significantly higher (M = .05) than those produced by younger adults (M = .00). The ANOVA revealed a significant interaction between segment and age ($F$[1, 17] = 6.42, $\eta_p^2$ = .27, $p < .05$) and the factors session and age ($F$[3, 51] = 10.02, $\eta_p^2$ = .37, $p$

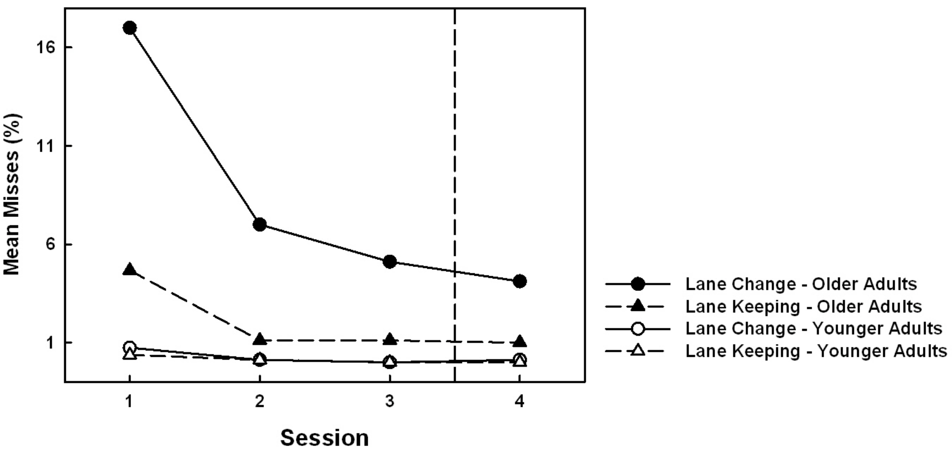


Figure C.17: Proportion of misses as a function of session (1-4), segment (lane change, lane keeping) and age group (younger, older).

$< .001$) as well as a significant 2-way interaction between session and segment ($F$[3, 51] = 4.91, $\eta_p^2 = .22$, $p < .05$). All those 2-way interactions were however modulated by a significant 3-way interaction including those factors (session x segment x age) ($F$[3, 51] = 3.55, $\eta_p^2 = .17$, $p < .05$). To further analyze this interaction, separate 2-way (session x segment) ANOVAs were conducted for younger and older adults respectively. There was no significant 2-way interaction between the factors segment and session for younger adults ($p > .55$), but the interaction reached significance for older adults ($F$[3, 27] = 5.26, $\eta_p^2 = .37$, $p < .05$). A follow-up of this interaction with two separate one-way (session) ANOVAs for older adults in LC and LK segments respectively, showed a significant simple main effect of session for LC segments ($F$[3, 27] = 15.99, $\eta_p^2 = .64$, $p < .001$) as well as LK segments ($F$[3, 27] = 8.31, $\eta_p^2 = .48$, $p < .01$). The interaction can be explained by practice having a stronger effect in LC segments as compared to LK segments. To further analyze these learning effects for both driving segments, paired-samples $t$-tests were performed between the first and third session. They revealed indeed that learning took place in both LC segments ($t$[19] = 4.06, $p < .01$) as well as LK segments ($t$[19] = 3.65, $p < .01$). The retention interval (analyzed with paired-samples $t$-tests between the third and fourth session) showed no significant effect for both LC ($p > .77$) as well as LK segments ($p > .74$). This means that the number of misses remained the same despite a practice-free period. The omnibus ANOVA revealed no other significant interactions including the factor segment (all $p$s $> .36$).

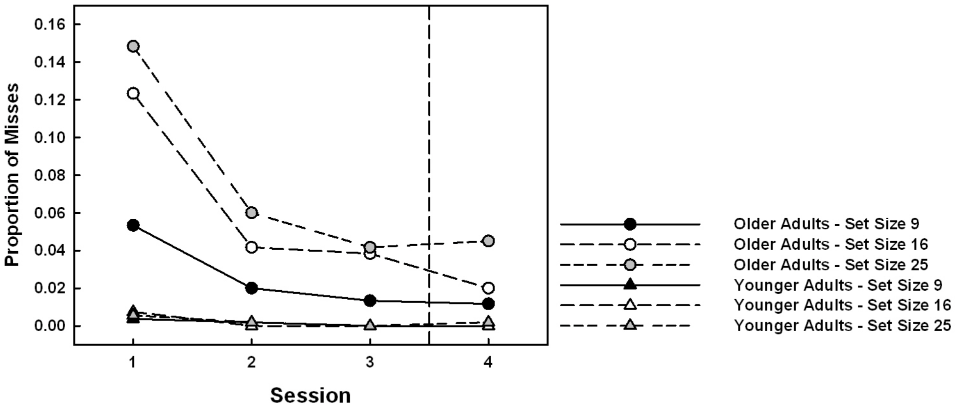


Figure C.18: Proportion of misses as a function of session (1-4), age group (younger, older) and set size (9, 16, 25).

Figure C.18 presents the proportion of misses as a function of session, age group and set size. The main effect of set size was significant ($F$[2, 34] $= 12.26$, $\eta_p^2 = .42$, $p < .001$): the search display with the highest set size provoked the highest proportion of misses (M $= .04$), followed by the medium set size (M $= .03$) and the smallest set size (M $= .01$). Set size interacted significantly with age as well ($F$[2, 34] $= 10.45$, $\eta_p^2 = .38$, $p < .01$). Two separate 1-way (set size) ANOVAs for each age group showed that there was a significant simple main effect of set size for older adults ($F$[2, 18] $= 13.52$, $\eta_p^2 = .60$, $p < .01$) but not for younger adults ($p > .64$). An increase in set size thus increased the proportion of misses for older adults, whereas for younger adults the effect of set size had no influence on the proportion of misses. Session interacted significantly with set size ($F$[6, 102] $= 3.48$, $\eta_p^2 = .17$, $p < .01$) as well. Three separate 1-way (session) ANOVAs for the set sizes 9, 16 and 25 respectively showed that there was a significant simple main effect of session for set size 9 ($F$[3, 54] $= 5.90$, $\eta_p^2 = .25$, $p < .05$), as well as for set size 16 ($F$[3, 54] $= 10.80$, $\eta_p^2 = .38$, $p < .001$) and finally set size 25 ($F$[3, 54] $= 13.34$, $\eta_p^2 = .43$, $p < .001$). The interaction comes from the differences in learning rate with practice: learning was strongest for the VST-display with 25 items, followed by the display with 16 items and finally the display with 9 items. The 3-way interaction including the factors session, set size and age showed a tendency, but just missed to reach significance ($F[6,\ 102] = 2.28$, $\eta_p^2 = .12$, $p = .06$). It indicates however that practice had hardly any effet on younger adults, whereas for older adults practice decreased the proportion of misses, especially for larger set sizes. All other interactions did not reach significance (all $p$s $> .26$).

### C.3.3 Dual-Task Measures

A separate 3-way ANOVA including the between-participants factor *age* (young, old) as well as the within-participants factors *session* (1-4) and *set size* (9, 16, 25) was conducted on the measure IRI. Eye behavior was analyzed with separate 4-way ANOVAs for the measures Vertical Saccade Frequency (VSF), Mean Glance Duration (MGD) and Total Glance Duration (TGD). These ANOVAs included the between-participants factor *age* (young, old) to account for age effects as well as three within-participants factors: *session* (1-4) to account for practice effects over several sessions, *segment* (lane change, lane keeping) to account for driving task complexity and finally the factor *set size* to account for the effect of secondary task complexity on eye behavior.

In total, 10057 glances were analyzed of which 4562 for lane change segments and 5495 for lane keeping segments. Glances were limited to the dual-task condition (as this is the only condition in which participants looked away from the road and towards the display and thus vertical saccades could be clearly detected) and to trials in which a correct response was given on the visual search task.

*IRI.* The ANOVA revealed a significant main effect of session ($F[3,\ 51] = 4.56$, $\eta_p^2 = .21$, $p < .05$). As can be observed in Figure C.19, with practice IRI decreased. A follow-up analysis with paired-samples $t$-tests for the first and third session showed a learning effect ($t[19] = 2.65$, $p < .05$): IRIs significantly decreased with practice. The analysis between the third and fourth session however showed that the retention interval had no effect on IRI data ($p > .20$). The ANOVA revealed a main effect of set size as well ($F[2,\ 34] = 42.96$, $\eta_p^2 = .72$, $p < .001$): IRI increased with increasing set size. This indicates that indeed an increase in set size caused more dual-task interference.

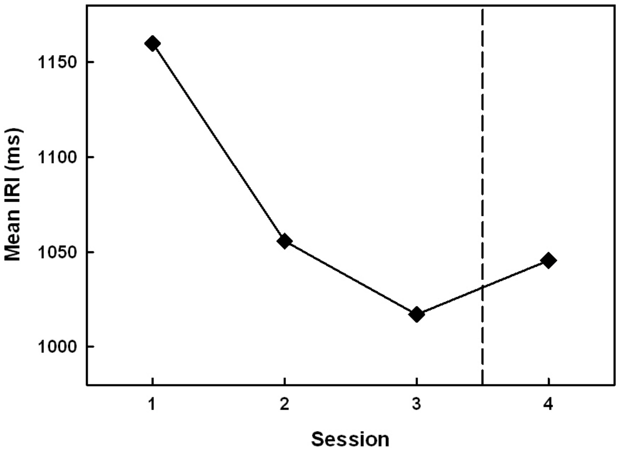


Figure C.19: Inter-Response Interval (IRI) as a function of session (1-4).

There was finally a main effect of age ($F$[1, 17] = 8.13, $\eta_p^2$ = .32, $p < .05$) with IRI of older adults being higher (M = 1219 ms) than that of younger adults (M = 921 ms), indicating that they "suffered" more from a secondary task while driving than younger adults. No significant interactions could be found (all $p$s > .33).

*Vertical Saccade Frequency.* Figure C.20 presents VSF as a function of session, age and segment. The ANOVA yielded a main effect of age ($F$[1, 18] = 14.194, $\eta_p^2$ = .44, $p < .01$) with older adults looking down at the search display less often (M = 18) than younger adults (M = 24). There was furthermore a main effect of segment ($F$[1, 108] = 12.051, $\eta_p^2$ = .40, $p < .01$): VSF were higher in LK segments (M = 23) as compared to LC segments (M = 19). Finally the ANOVA revealed a main effect of set size ($F$[1, 108] = 55.56, $\eta_p^2$ = .76, $p < .001$): the smallest set sizes yielded the highest number of display glances ($M_{\text{Setsize9}}$ = 23) and they decreased with an increase in set size ($M_{\text{Setsize16}}$ = 21; $M_{\text{Setsize25}}$ = 19). There was no main effect of session, which indicates that the number of display glances did not change with practice. The interaction between session and age was significant ($F$[1, 108] = 5.07, $\eta_p^2$ = .22, $p < .05$). To follow up on this ANOVA we conducted two separate 1-way (session) ANOVAs for younger and older adults respectively. The factor session had no effect on younger adults ($p > .66$) indicating that their vertical saccade behavior did not change with practice. However, there was a significant

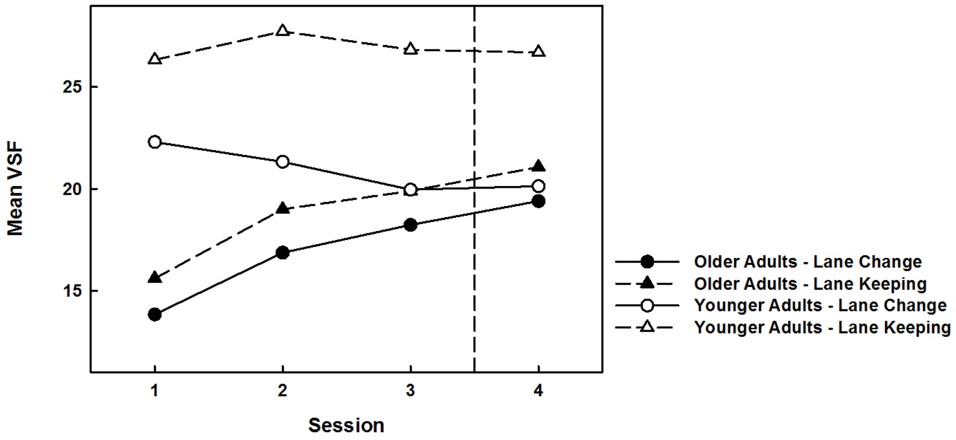


Figure C.20: Mean vertical saccade frequency (VSF) as a function of session (1-4), age group (younger, older) and segment (lane change, lane keeping).

simple main effect of session for older adults ($F[3, 27] = 9.59$, $\eta_p^2 = .52$, $p < .01$), which indicated that with practice, VSF for older adults increased. This might reflect that with practice, older adults feel more confident while performing the task and take their eyes more off the road to look at the display more often. A follow-up analysis of this learning effect for older adults with paired-samples $t$-tests showed that there was indeed a learning effect which took place between the first and third session ($t[9] = 2.93$, $p < .05$). The retention interval had no effect on VSF: there was no significant difference in VSF between the third and fourth session ($p > .18$). The interaction between session and segment showed a tendency, but did not reach significance ($F[1, 108] = 3.45$, $\eta_p^2 = .16$, $p = .08$). It seems to indicate though that with practice, VSF in the LK segments increases more than in the LC segments. Finally there was a significant interaction between age and set size ($F[1, 108] = 17.31$, $\eta_p^2 = .49$, $p < .001$) which is represented in Figure C.21. To follow up on this interaction two 1-way (set size) ANOVAs were conducted for each age group respectively. There was a significant simple main effect of set size for younger adults ($F[2, 18] = 7.811$, $\eta_p^2 = .46$, $p < .01$) as well as for older adults ($F[2, 18] = 68.38$, $\eta_p^2 = .88$, $p < .001$). With increasing set size, VSF towards the secondary task display decreased for both age groups, whereby the decrease was stronger for older adults as compared to younger adults.

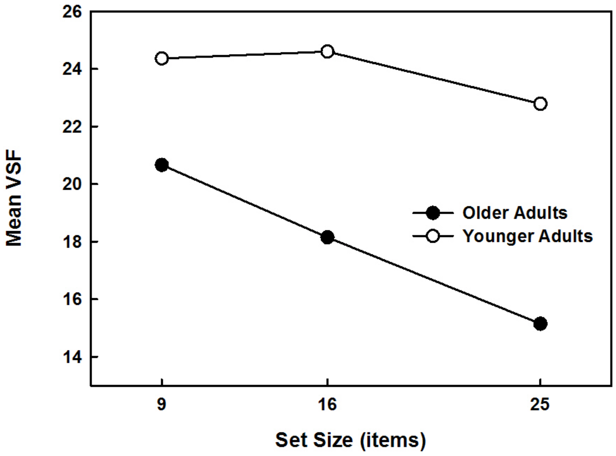


Figure C.21: Mean Vertical Saccade Frequency (VSF) as a function of set size (9, 16, 25) and age group (younger, older).

*Glance Duration.* Figure C.22 presents mean MGD as a function of session, age group and segment. The ANOVA yielded a main effect of age ($F$[1, 18] = 38.38, $\eta_p^2 = .68$, $p < .001$) with MGD of older adults (M = 1041 ms) being significantly longer than those of younger adults (M = 821 ms). The ANOVA furthermore revealed a main effect of segment ($F$[1, 108] = 4.93, $\eta_p^2 = .22$, $p < .05$): MGD towards the secondary task display were significantly lower in LC segments (M = 904 ms) as compared to LK segments (M = 958 ms). This indicates that in the more difficult driving conditions, participants took their eyes less off the road. Finally, as can be seen in figure C.23, which presents MGD as a function of age group and set size, we found a main effect of set size ($F$[1, 108] = 107.92, $\eta_p^2 = .86$, $p < .001$): larger set sizes provoked the longest glance durations ($\mathrm{M}_{\mathrm{Setsize25}}$ = 1029 ms; $\mathrm{M}_{\mathrm{Setsize16}}$ = 936 ms; $\mathrm{M}_{\mathrm{Setsize9}}$ = 828 ms). No other main effects or interactions were found (all $p$s > .31).

*Total Glance Duration.* Figure **??**a presents mean TGD as a function of session and age group. Figure **??**b presents mean TGD as a function of set size and age group. The ANOVA yielded a main effect of segment ($F$[1, 18] = 19.38, $\eta_p^2 = .52$, $p < .001$): TGD was lower in LC segments (M = 21248 ms) than in LK segments (M = 16908 ms). There

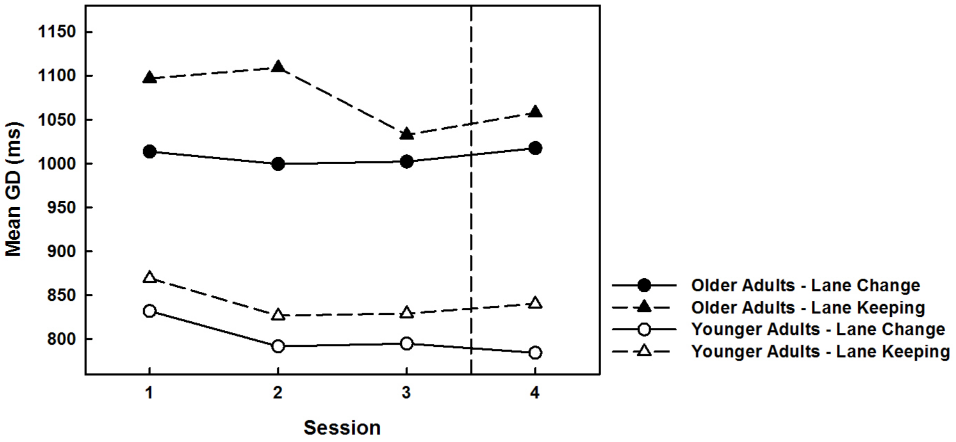


Figure C.22: Mean Glance Duration (MGD) as a function of set size (9, 16, 25) and age group (younger, older).

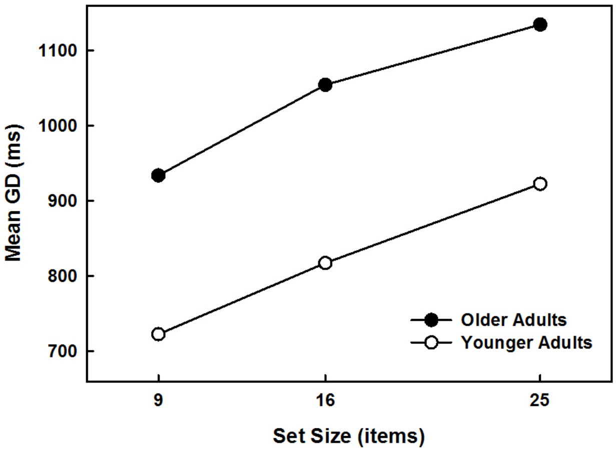
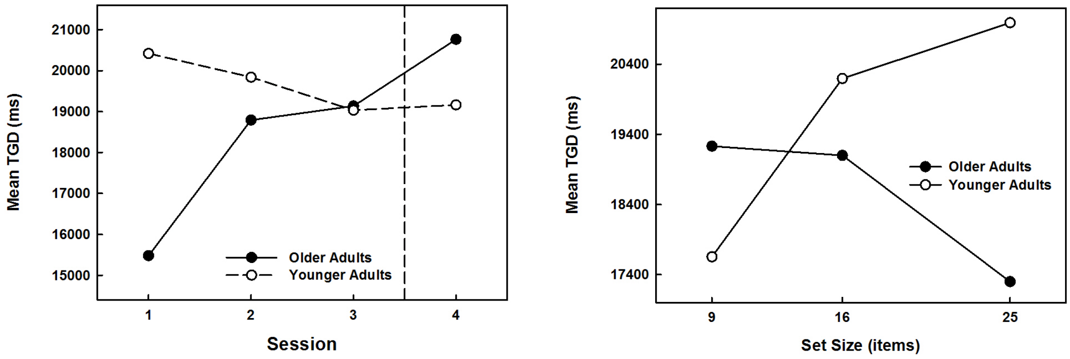


Figure C.23: Mean Glance Duration (MGD) as a function of set size (9, 16, 25) and age group (younger, older).

was also a main effect of set size ($F$[2, 36] = 3.62, $\eta_p^2$ = .17, p < .05) with TGD being lowest in the visual search task with the smallest set size of 9 items (M = 18442 ms), followed by the visual search task with the highest set size of 25 items (M = 19142 ms) and the longest TGD for the medium set size of 16 items (19647 ms). There was furthermore a significant interaction between session and age group ($F$[1, 108] = 4.74, $\eta_p^2$ = .21, $p <$ .05). To follow up on this interaction separate 1-way (session) ANOVAs were conducted for younger and older adults respectively. There was no significant siomple main effect of session on TGD for younger adults ($p > .55$), but there was a significant simple main effect of session for older adults ($F$[3, 27] = 6.38, $\eta_p^2$ = .42, $p < .05$) indicating that the TGD of older adults increased significantly with practice. A follow-up analysis on this effect of session with paired-samples $t$-tests showed that the simple main effect actually came from a significant learning effect between the first and second session only ($t$[9] = 2.61, $p < .05$). The comparison between the first and the third session showed a tendency ($t$[9] = 2.13, $p = .06$), but did not reach significance. It indicates however that adding an extra session has a potential effect on TGD. The comparison between the third an fourth session, taking into account the retention period, showed a tendency as well, but missed to reach significance ($t$[9] = 2.17, $p = .06$). It might however indicate that even without practicing the task, some learning might take place and therewith influencing the TGD.

The omnibus ANOVA finally revealed a significant interaction between age and set size ($F$[1, 108] = 17.16, $\eta_p^2$ = .49, $p < .001$) as shown in Figure **??**b. To follow up on this interaction two separate 1-way (set size) ANOVAs were conducted for each age group. There was a significant simple main effect of set size on TGD of younger adults ($F$[2, 18] = 24.51, $\eta_p^2$ = .73, $p < .001$) as well as for older adults ($F$[2, 18] = 4.15, $\eta_p^2$ = .32, $p$ < .05). An increase in set size had opposite effects on the different age groups though: for younger adults, TGD towards the secondary task display increased with increasing set size, whereas for older adults, TGD towards the secondary task display decreased with increasing set size. This might be an indication that older adults actually "suffer more" from an increasing set size than younger adults (who seem to be more comfortable with the task at hand, allowing them to take their eyes off the road longer when the secondary task gets harder).



(a) (b)

Figure C.24: Total Glance Duration (TGD) as a function of (a) session (1-4) and age group (young, old) and as a function of (b) set size (9, 16, 25) and age group (younger, older).

# D Appendix 4: Results Experiment 4

Results are divided in two sections: First we analyzed the effect that adding a (relevant or non-relevant) secondary task had on performance measures by comparing single- to dual-task data. Then we analyzed the dual-task conditions "dual" and "dual-r" more in detail. For both sections, we examined driving, visual-search, dual-task, and subjective data separately. The different ANOVAs used for each section are described at the beginning of the section concerned. If a significant main effect or a significant simple main effect with the factor *session* was present, power functions were fitted to quantify how performance changes with learning. The equation used for fitting was the following:

$$y = a + b * x^{-c}$$

where $y$ represents a dependent variable, $x$ represents the session number, $a$ represents the asymptote, $b$ is the amount of time required at the beginning of training and $c$ specifies the rate of speed-up with training. $R^2$ and Root Mean Square Error (RMSE) were used to quantify goodness of fit. Arcsine-transformed data for which power functions needed to be fitted, are always represented with two graphs: One representing the non-arcsine transformed data (as this is what is typically presented) and another one representing the arcsine-transformed data including the fitted curves (as these are the actual data the ANOVAs were performed on).

Due to the amount of data, figures were only included for significant or close to significant ($p_{max}$ =.08) main effects or interactions. Note that close to significant interactions will only be described, without providing any exploratory follow-up tests.

# D.1 Comparing Task Types: Single-Task, Dual-Task and Dual-R

## D.1.1 Driving Measures

Separate 3-way mixed-factors ANOVAs were conducted on the driving measures RT-LC, MT-LC, MDEV, SDDEV and SDSW. The ANOVA included the within-participants factor *session* (1-4) to analyze the effect of practice over sessions, the within-participants factor *task type* (single, dual, dual-r) to look at differences between task types and finally the between-participants factor *age* (younger, older) to account for age effects. All measures were based on LC segments only and all ANOVAs were averaged over the factor set size.

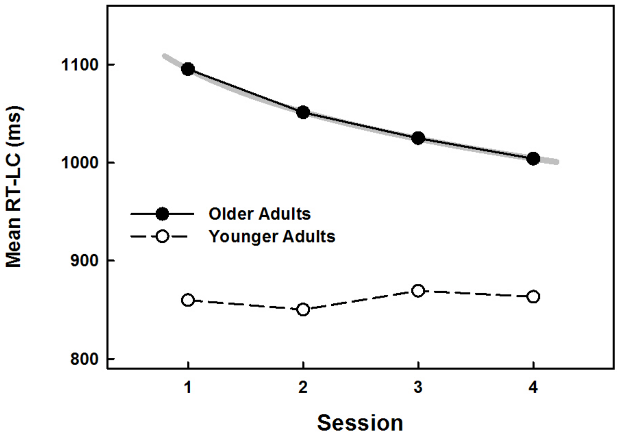


Figure D.1: Mean Reaction Time until Lane Change (RT-LC) as a function of session (1-4) and age (younger, older) with fitted power function for the older age group.

*RT-LC.* Figure D.1 presents mean RT-LC as a function of session and age with fitted power function for the older age group. The ANOVA revealed a significant main effect of task type on RT-LC ($F[2, 58] = 9.52$, $\eta_p^2 = .53$, $p < .001$) with mean RT-LC being highest in the dual-r condition (M = 994 ms) followed by the dual-condition (M = 954 ms) and finally the LCT-single condition (M = 907 ms). There was furthermore a significant main effect of session on RT-LC ($F[3, 87] = 3.97$, $\eta_p^2 = .12$, $p < .05$): RT-LC decreased over sessions. There was a main significant effect of age as well ($F[1, 29] = 14.18$, $\eta_p^2 = .98$, $p < .001$): RT-LC was significantly higher for older adults (M = 1043) as compared to younger adults (M = 860 ms).

The interaction between session and age was significant ($F[3, 87] = 5.49$, $\eta_p^2 = .16$, $p < .01$). Follow-up tests with separate 1-way (session) ANOVAs for older and younger adults respectively, revealed a simple main effect of session for older adults ($F[3, 45] = 6.58$, $\eta_p^2 = .31$, $p < .01$; fitted power function: $y = 1604.32 - 509.18x^{-0.12}$, $R^2 = .99$, RMSE = 2.30) but not for younger adults ($p > .55$). This means that with practice, RT-LC values decreased significantly for older adults, but not for younger adults.

The omnibus ANOVA revealed no other significant interactions (all $p$s > .49).

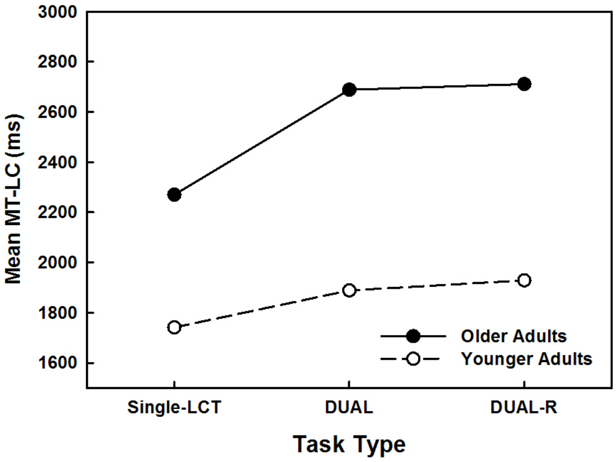


Figure D.2: Mean Movement Time (MT-LC) as a function of task type (single-LCT, dual, dual-r) and age (younger, older).

*MT-LC.* Figure D.2 presents mean MT-LC as a function of task type and age. The ANOVA revealed a significant main effect of task type on MT-LC ($F[2, 58] = 33.55$, $\eta_p^2 = .53$, $p < .001$) with mean MT-LC being highest in the dual-r condition (M = 2319 ms) followed by the dual-condition (M = 2288 ms) and finally the LCT-single condition (M = 2005 ms). There was a main effect of age as well ($F[1, 29] = 14.18$, $\eta_p^2 = .36$, $p < .001$) with MT-LCs for older adults being significantly higher (M = 2556 ms) than those of younger adults (M = 1852). The interaction between task type and age was significant ($F[2, 58] = 6.41$, $\eta_p^2 = .18$, $p < .05$). Follow-up tests with separate 1-way (task type) ANOVAs for older and younger adults respectively, revealed a simple main effect of task type for older adults ($F[2, 30] = 17.41$, $\eta_p^2 = .54$, $p < .001$) as well as for younger adults ($F[2, 30] = 32.70$, $\eta_p^2 = .69$, $p < .001$). The interaction can be explained by the fact that with increasing task-type difficulty, the MT-LC values for older adults increase more than those of younger adults. This seems to be the case especially when going from single-LCT to dual, but not from dual to dual-r.

The omnibus ANOVA revealed no other significant interactions (all $ps > .18$).

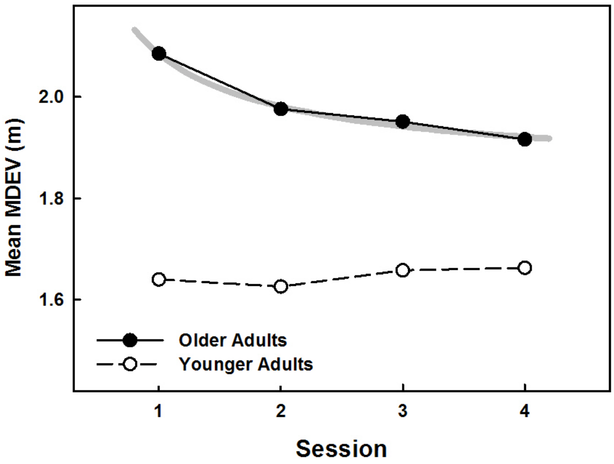


Figure D.3: Mean Lateral Deviation (MDEV) as a function of session (1-4) and age (younger, older) with fitted power function for the older age group.

*MDEV.* Figure D.3 presents mean MDEV as a function of session and age with fitted power function for the older age group. The ANOVA revealed a significant main effect of task type on MDEV ($F[2, 58] = 46.84$, $\eta_p^2 = .62$, $p < .001$) with mean MDEV being highest in the dual-r condition (M = 1.91 m) followed by the dual-condition (M = 1.82

m) and finally the LCT-single condition (M = 1.71 m). There was a main effect of session as well ($F[3, 87] = 3.52$, $\eta_p^2 = .11$, $p < .05$): Mean MDEV values decreased with practice. Finally there was a main significant effect of age ($F[1, 29] = 15.80$, $\eta_p^2 = .35$, $p < .001$) with mean MDEV being significantly higher for older adults (M = 1.98 m) as compared to younger adults (M = 1.65 m). The interaction between session and age was significant ($F[3, 87] = 5.64$, $\eta_p^2 = .16$, $p < .01$). A further analysis of this interaction with separate 1-way (session) ANOVAs for older and younger adults respectively, revealed a simple main effect of session for older adults ($F[3, 45] = 7.16$, $\eta_p^2 = .32$, $p <$ .01; fitted power function: $y = 1.84 + 0.25x^{-0.79}$, $R^2 = .99$, RMSE = .001) but not for younger adults ($p > .56$). This means that over sessions, lane-change accuracy increased significantly for older adults, but not for younger adults. Again, this observation might be due to the younger adults being at maximum performance from the first session on, leaving little to no room for improvement on their performance.

The omnibus ANOVA revealed no other significant main effects or interactions (all *p*s $> .15$).

*SDDEV.* The ANOVA revealed a significant main effect of task type on SDDEV ($F[2, 58] = 32.28$, $\eta_p^2 = .53$, $p < .001$) with mean SDDEV being highest in the dual-r condition (M = 1.53 m) followed by the dual-condition (M = 1.45 m) and finally the LCT-single condition (M = 1.42 m). There was furthermore a significant main effect of age on SDDEV as well ($F[1, 29] = 8.68$, $\eta_p^2 = .36$, $p < .001$) with mean SDDEV being higher for older adults (M = 1.55 m) as compared to younger adults (M = 1.39 m). No other significant main effects or interactions could be found (all *p*s $> .17$).

*SDSW.* Figure D.4 presents mean SDSW as a function of session. The ANOVA revealed a significant main effect of task type on SDSW ($F[2, 58] = 37.52$, $\eta_p^2 = .56$, $p < .001$) with mean SDSW being lowest in the dual-r condition (M = 12.45°) followed by the dual-condition (M = 12.89°) and finally the LCT-single condition (M = 15.67°). There was furthermore a main significant effect of age on SDSW ($F[1, 29] = 17.49$, $\eta_p^2 = .38$, $p < .001$) with mean SDSW values being higher for younger adults (M = 17.30°) as compared to older adults (10.04°). The main effect of session just missed to reach significance ($p = .06$), but seems to provide a hint that with practice mean SDSW val-

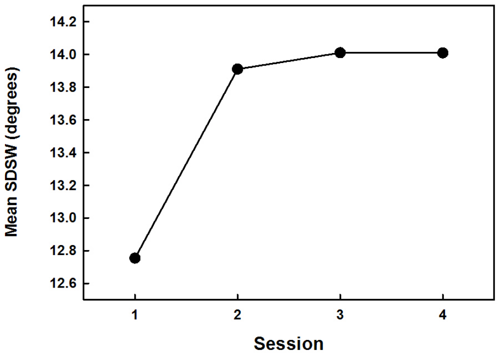


Figure D.4: Mean Standard Deviation of Steering Wheel Angle (SDSW) as a function of session (1-4).

ues increase, leading to more corrective steering-wheel movements. No other significant interactions could be found (all $p$s $>$ .35).

### D.1.2 Visual Search Measures

Performance on the VST was analyzed by the RT-VST as well as computing the proportion of incorrect responses and the proportion of misses for each condition and participant. Values for incorrect responses and misses were arcsine transformed to deal with the non-normality of proportions (e.g., see Winer, 1971). Separate 4-way mixed-factors ANOVAs were conducted on the visual-search measures RT-VST and the proportion of incorrect responses. All ANOVAs included the between-participants factor *age* (younger, older) to establish the effect of age, the within-participants factor *task type* (single-VST, dual, dual-r) to examine the effect of task complexity, the effect of practice with the factor *session* (1-4) and finally the effect of secondary-task complexity with the factor *set size* (9, 25).

*RT-VST*. Figure D.5 presents mean RT-VST as a function of session, age and set size with fitted power functions for all conditions. The ANOVA revealed a significant main effect of task type on RT-VST ($F[2, 64] = 39.91$, $\eta_p^2 = .56$, $p < .001$). Reaction times on the VST were highest in the dual-r condition (M = 1187 ms), followed by the

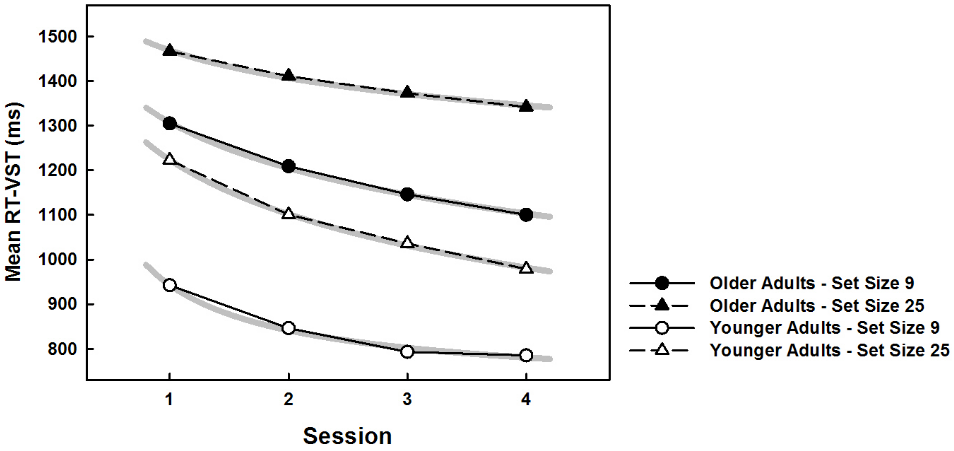


Figure D.5: Mean Reaction Time on the Visual Search Task (RT-VST) as a function of session (1-4), age (younger, older) and set size (9, 25) with fitted power functions for all conditions.

dual condition (M = 1152 ms) and finally the single-VST condition (M = 1047 ms). The factor session showed a significant main effect as well ($F$[3, 96] = 100.28, $\eta_p^2 = .76$, $p < .001$): With practice, RT-VST significantly decreased. There was furthermore a significant main effect of set size ($F$[1, 32] = 279.44, $\eta_p^2 = .90$, $p < .001$) with RT-VST values being significantly higher for the 25-items set size (M = 1241 ms) as compared to the 9 items set size (M = 1016 ms). Finally there was a main effect of age ($F$[1, 32] = 101.17, $\eta_p^2 = .76$, $p < .001$): RT-VSTs were higher for older adults (M = 1294 ms) as compared to younger adults (M = 963 ms). The interaction between task type and age showed a tendency, but missed to reach significance ($p = .09$). This might be an indication however that older adults suffer more from an increase in task complexity as compared to younger adults.

The 3-way interaction between session, set size and age shown in Figure D.5 was significant ($F$[3, 96] = 15.31, $\eta_p^2 = .32$, $p < .001$). A follow-up of this interaction with two separate 2-way (session x set size) ANOVAs for younger and older adults respectively, revealed a significant interaction between both factors for younger adults ($F$[3, 48] = 8.91, $\eta_p^2 = .36$, $p < .001$) as well as older adults ($F$[3, 48] = 4.48, $\eta_p^2 = .22$, $p < .05$).

Follow-up analyses for the younger age group with separate 1-way (session) ANOVAs for each set size respectively, revealed a simple main effect of session both for the 9-item display ($F$[3, 48] = 38.29, $\eta_p^2$ = .71, $p$ < .001; fitted power function: $y$ = 689 + 253$x$ $^{-0.74}$, R$^2$ = .99, RMSE = 11.35) as well as the 25-item display ($F$[3, 48] = 47.69, $\eta_p^2$ = .75, $p$ < .001; fitted power function: $y$ = -7255 + 8478$x$ $^{-0.02}$, R$^2$ = .99, RMSE = 5.89). Follow-up analysis for the older age group with separate 1-way (session) ANOVAs for set size 9 and set size 25 respectively, revealed a simple main effect of session for both the 9-item display ($F$[3, 48] = 33.47, $\eta_p^2$ = .68, $p$ < .001; fitted power function: $y$ = -13759 + 15066$x$ $^{-0.01}$, R$^2$ = .99, RMSE = 5.86) as well as the 25-item display ($F$[3, 48] = 7.59, $\eta_p^2$ = .32, $p$ < .001; fitted power function: $y$ = -12219 + 13688$x$ $^{-0.01}$, R$^2$ = .99, RMSE = 6.10). The 3-way interaction shown in Figure D.5 can thus be explained as follows: With practice mean RT-VST values decrease for both younger and older adults for all set sizes. However, as shown by the fitted functions, the speed at which learning takes place differs between age group and set size (as indicated by the rates of speed-up with practice, indicated in the power functions). Learning is fastest for the younger adults and the smallest set size. Learning gets slower for the younger adults and the bigger set size. Finally, learning is slowest for the oldest adults for both set sizes.

The omnibus ANOVA revealed no other significant interactions (all $p$s > .20).

*Incorrect Responses.* Figure D.6 presents the mean proportion of incorrect responses on the VST as a function of task type and age, with fitted power functions for the older adults, set size 9 and 25 and the younger adults, set size 25. Figure **??** presents the mean proportion of incorrect responses on the VST as a function of session and age for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions. Finally, Figure **??** presents the mean proportion of incorrect responses as a function of session, age and set size for (a) non-arcsine transformed and (b) arcsine-transformed values. The ANOVA yielded a significant main effect of task type on the number of incorrect responses on the VST ($F$[2, 64] = 22.72, $\eta_p^2$ = .42, $p$ < .001): The highest proportion of incorrect responses were given in the dual-r condition (M = .13), followed by the dual condition (M = .13) and finally the single-VST condition (M = .09). There was furthermore a significant main effect of session ($F$[3, 96] = 56.50, $\eta_p^2$ = .64, $p$ < .001): With practice, the proportion of incorrect responses decreased significantly.

There was a significant main effect of set size as well ($F[1,\ 32] = 404.10$, $\eta_p^2 = .93$, $p < .001$): The proportion of incorrect responses was significantly lower for the 9-item displays (M = .04) as compared to the 25-items displays (M = .19). Finally there was a significant main effect of age ($F[1,\ 32] = 114.68$, $\eta_p^2 = .78$, $p < .001$) with older adults producing more incorrect responses (M = .20) than younger adults (M = .04).

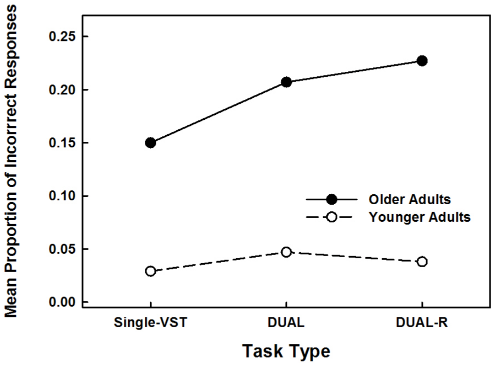
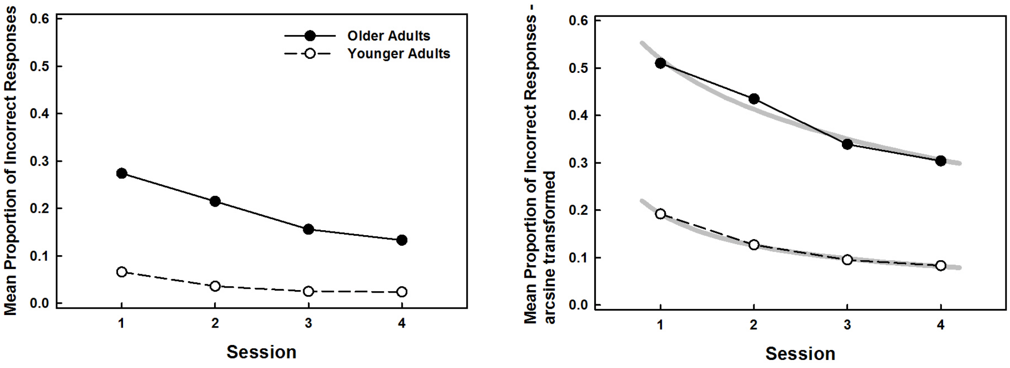


Figure D.6: Mean Proportion of Incorrect Responses on the Visual Search Task as a function of task type (single-VST, dual, dual-r) and age (younger, older).

The ANOVA revealed a significant 2-way interaction for the factors task type and age ($F[2,\ 64] = 7.48$, $\eta_p^2 = .19$, $p < .001$). In order to follow-up on this interaction 2 separate 1-way (task type) ANOVAs were performed for each age group. There was a significant simple main effect of session for the older adults ($F[2,\ 32] = 15.23$, $\eta_p^2 = .49$, $p < .001$), but not for the younger adults ($p > .12$). The interaction shown in Figure D.6 can thus be explained by the fact that with increasing overall task-difficulty, the proportion of incorrect responses does not increase statistically for younger adults, but increase for older adults.

The ANOVA furthermore revealed a significant interaction between the factors session and age ($F[3,\ 96] = 6.44$, $\eta_p^2 = .17$, $p < .01$) as shown in Figure D.7. Follow-up analyses with separate 1-way (session) ANOVAs for each age group, revealed a simple main effect of session for younger adults ($F[3,\ 48] = 11.21$, $\eta_p^2 = .41$, $p < .01$; fitted power function: $y = \text{-}0.00 + 0.20x^{\text{-}0.60}$, $R^2 = .99$, RMSE = 0.00) as well as for older adults ($F[3,\ 48] = 44.05$, $\eta_p^2 = .73$, $p < .001$; fitted power function: $y = \text{-}229.15 + 229.67x^{\text{-}0.000668099}$,

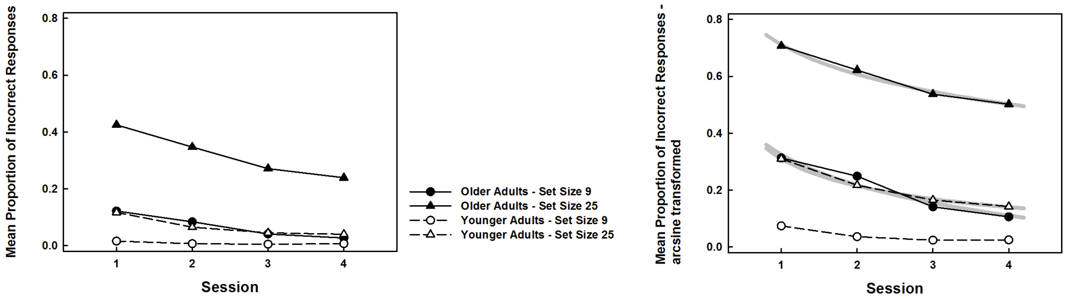


(a) (b)

Figure D.7: Proportion of Incorrect Responses as a function of session (1-4) and age (younger, older) for (a) non-arcsine transformed and (b) arcsine-transformed values with fitted power functions.

$R^2$ = .97, RMSE = 0.03). The interaction can thus be explained by the fact that with practice, the proportion of errors reduces faster for younger adults as compared to older adults.

The ANOVA revealed a significant interaction between set size and age ($F$[1, 32] = 63.30, $\eta_p^2$ = .66, $p$ < .001) as well as the factors session and set size ($F$[3, 96] = 2.92, $\eta_p^2$ = .08, $p$ < .05). Those interactions were however modulated by a significant 3-way interaction (shown in Figure D.8) including the factors session, age and set size ($F$[3, 96] = 3.66, $\eta_p^2$ = .10, $p$ < .05). A follow-up analysis with separate 2-way (session x set size) ANOVAs for each age group respectively, revealed a significant interaction for both factors for younger adults ($F$[3, 48] = 14.94, $\eta_p^2$ = .48, $p$ < .001) as well as older adults ($F$[3, 48] = 4.87, $\eta_p^2$ = .23, $p$ < .05). Follow-up analyses with separate 1-way (session) ANOVAs for each set size and each age group, revealed a significant simple main effect for younger adults, set size 25 ($F$[3, 48] = 13.70, $\eta_p^2$ = .46, $p$ < .001; fitted power function: $y$ = -0.12 + 0.43$x^{\text{-0.36}}$, $R^2$ = .99, RMSE = 0.01), as well as for older adults, set size 9 ($F$[3, 48] = 23.63, $\eta_p^2$ = .60, $p$ < .001; fitted power function: $y$ = -305.07 + 305.39$x^{\text{-0.00}}$, $R^2$ = .95, RMSE = 0.04) and set size 25 ($F$[3, 48] = 26.00, $\eta_p^2$ = .62, $p$ < .001; fitted power function: $y$ = -134.70 + 135.41$x^{\text{-0.00}}$, $R^2$ = .99, RMSE



(a) (b)

Figure D.8: Proportion of Incorrect Responses as a function of session (1-4), age (younger, older) and set size (9, 25) for (a) non-arcsine transformed values and (b) arcsine-transformed values.

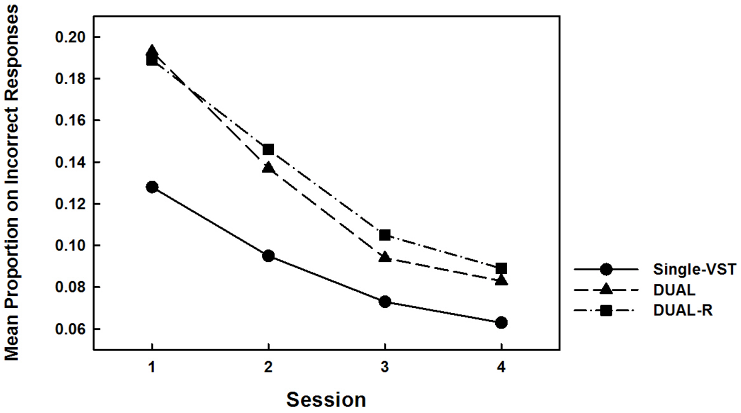
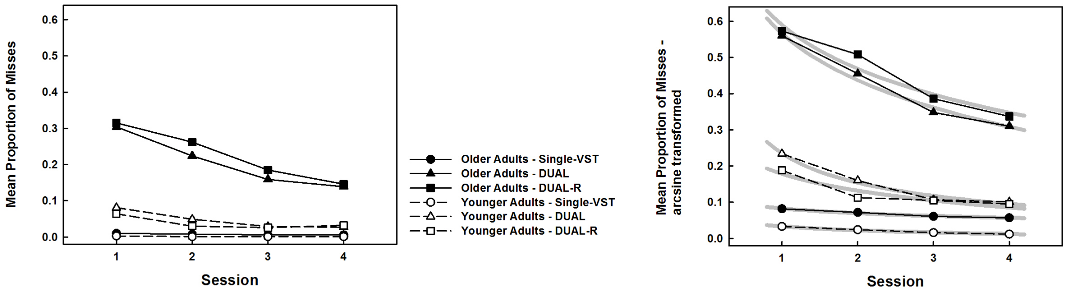


Figure D.9: Mean Proportion of Incorrect Responses on the Visual Search Task as a function of session (1-4) and task type (single-VST, dual, dual-r) with fitted power functions for the older adults, set size 9 and 25 and the younger adults, set size 25.



(a) (b)

Figure D.10: Proportion of misses on the Visual Search Task as a function of session (1-4), age (younger, older) and task type (single-VST, dual, dual-r) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions.

= 0.02). The interaction shown in Figure D.8 can thus be explained by the fact that with practice, the proportion of incorrect responses decreases for younger adults with the bigger set size only, whereas for older adults, the proportion of incorrect responses decreases for both set sizes. The observed differences between learning effects for the different set sizes for older adults, comes from differences in learning rates as shown by the fitted power functions.

Figure D.9 presents the mean proportion of incorrect responses on the VST as a function of session and task type. The 2-way interaction between both factors showed a tendency, but missed to reach significance ($F[6, 192] = 1.98$, $\eta_p^2 = .06$, $p = .08$). It might provide an indication however that with practice, the proportion of incorrect responses decreases and this decrease seems to be stronger for the dual- and dual-r condition as compared to the single-VST condition.

The omnibus ANOVA revealed no other significant interactions (all $p$s $> .19$).

*Misses.* Figure D.10 presents the mean proportion of misses on the VST as a function of session, age and task type for (a) non-arcsine transformed and (b) arcsine-transformed values, with fitted power functions. Figure **??** presents the mean proportion of misses on the VST as a function of task type, age and set size. Finally, Figure D.12 presents the mean proportion of misses on the VST as a function of session, age and set size for (a) non-arcsine transformed and (b) arcsine-transformed values.

The ANOVA yielded a significant main effect of task type on the number of misses on the VST ($F[2, 64] = 198.01$, $\eta_p^2 = .86$, $p < .001$): The highest proportion of misses were observed in the dual-r condition (M = .13), followed by the dual condition (M = .13) and finally the single-VST condition (M = .00). There was furthermore a significant main effect of session ($F[3, 96] = 57.50$, $\eta_p^2 = .64$, $p < .001$): With practice, the proportion of misses on the VST decreased significantly. There was a significant main effect of set size as well ($F[1, 32] = 348.82$, $\eta_p^2 = .92$, $p < .001$): The proportion of misses was significantly higher for the 25-item displays (M = .14) as compared to the 9-items displays (M = .03). Finally there was a significant main effect of age ($F[1, 32] = 101.33$, $\eta_p^2 = .76$, $p < .001$) with older adults yielding more misses (M = .15) than younger adults (M = .03).

There was a significant 2-way interaction between task type and age ($F[2, 64] = 55.18$, $\eta_p^2 = .63$, $p < .001$), session and age ($F[3, 96] = 8.26$, $\eta_p^2 = .21$, $p < .001$) and task type and session ($F[6, 192] = 15.98$, $\eta_p^2 = .33$, $p < .001$). Those interactions were however modulated by a significant 3-way interaction including the factors session, task type and age ($F[6, 192] = 3.53$, $\eta_p^2 = .01$, $p < .01$) presented in Figure D.10. A follow-up analysis of this interaction with separate 2-way (session x task type) ANOVAs for each age group, revealed a significant interaction between those factors for older adults ($F[6, 96] = 19.50$, $\eta_p^2 = .25$, $p < .01$) as well as younger adults ($F[6, 96] = 3.04$, $\eta_p^2 = .55$, $p < .001$). Follow-up tests of those interactions with separate 1-way (session) ANOVAs for each task type and each age group respectively, revealed a simple main effect of session for the single-VST ($F[3, 48] = 5.33$, $\eta_p^2 = .25$, $p < .01$; fitted power function: $y = -12.67 + 12.70x^{-0.00121264}$, $R^2 = .99$, RMSE = 0.00), the dual- ($F[3, 48] = 10.77$, $\eta_p^2 = .40$, $p < .001$; fitted power function: $y = -0.05 + 0.29x^{-0.48}$, $R^2 = .98$, RMSE = 0.01) and the dual-r condition ($F[3, 48] = 3.50$, $\eta_p^2 = .18$, $p < .05$; fitted power function: $y = 389.98 - 389.80x^{-0.0002}$, $R^2 = .90$, RMSE = 1.48) for younger adults and a simple main effect of session for single-VST ($F[3, 48] = 5.26$, $\eta_p^2 = .25$, $p < .01$; fitted power function: $y = -17.68 + 17.76x^{-0.00104807}$, $R^2 = .98$, RMSE = 0.00), dual ($F[3, 48] = 38.02$, $\eta_p^2 = .70$, $p < .001$; fitted power function: $y = -149.31 + 149.88x^{-0.00124267}$, $R^2 = .99$, RMSE

= 0.02) and dual-r ($F$[3, 48] = 45.48, $\eta_p^2$ = .74, $p < .001$; fitted power function: $y$ = -437.68 + 438.27$x$ $^{-0.000399812}$, $R^2$ = .94, RMSE = 0.05) conditions for older adults. The 3-way interaction shown in Figure D.10 can thus be explained as follows: With practice, the proportion of misses reduces significantly for both younger and older adults and for all task-type conditions. As can be seen in Figure D.10, the reduction in proportion of misses is highest for older adults in the dual- and dual-r conditions. The proportion of misses for the dual- and dual-r condition for younger adults reduces as well, but not as strong as for older adults. And although graphically, the proportion of misses for VST for both younger and older adults seems the same for each session, the ANOVAs show that there is actually some learning taking place, thereby decreasing the number of misses produced.

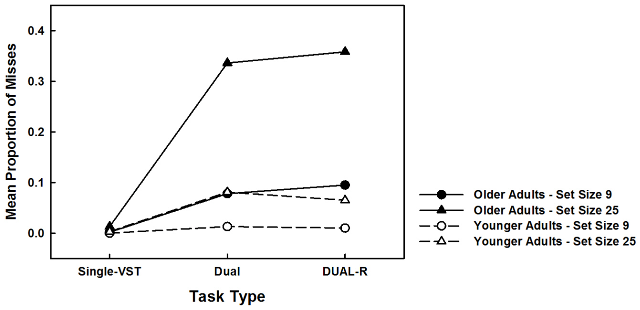
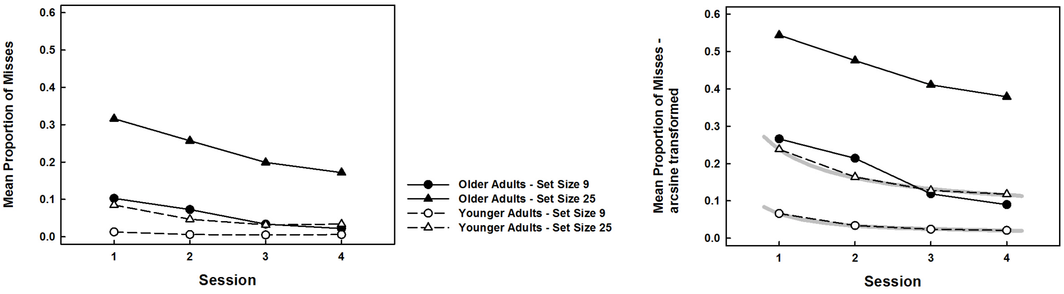


Figure D.11: Proportion of misses on the Visual Search Task as a function of task type (single-VST, dual, dual-r), age (younger, older) and set size (9, 25).

There was a significant 2-way interaction between the factors set size and age ($F$[1, 32] = 50.44, $\eta_p^2$ = .61, $p < .001$) as well as the factors task type and set size ($F$[2, 64] = 117.13, $\eta_p^2$ = .79, $p < .001$). Those interactions were, however, modulated by a significant 3-way interaction including the factors task type, set size and age ($F$[2, 64] = 15.05, $\eta_p^2$ = .32, $p < .001$). A follow-up analysis with separate 2-way (task type x set size) ANOVAs for each age group, revealed a significant interaction between both factors for younger adults ($F$[2, 32] = 25.07, $\eta_p^2$ = .61, $p < .001$) as well as older adults ($F$[2, 32] = 104.78, $\eta_p^2$ = .87, $p < .001$). A follow-up of these interactions with separate 1-way (task type) ANOVAs for each set size and each age group, revealed a simple main effect of session for set size 9 ($F$[2, 32] = 50.61, $\eta_p^2$ = .76, $p < .001$) and set size 25 ($F$[2,



(a) (b)

Figure D.12: Proportion of misses on the Visual Search Task as a function of session (1-4), age (younger, older) and set size (9, 25) for (a) non-arcsine transformed values and (b) arcsine-transformed values, with fitted power functions for the younger age group.

32] = 288.60, $\eta_p^2$ = .95, $p < .001$) for younger adults. For older adults, there was as well a simple main effect of session for set size 9 ($F$[2, 32] = 7.62, $\eta_p^2$ = .32, $p < .01$) and set size 25 ($F$[2, 32] = 38.57, $\eta_p^2$ = .71, $p < .001$). The 3-way interaction represented in Figure D.11 can thus be explained by the fact that with an overall increase in task difficulty, the proportion of misses increases for both age groups. However, as can be seen visually, the increase is stronger for the 25-items set size as compared to the 9-items set size. And although not graphically visible, there is actually a significant increase in the proportion of misses over sessions for younger adults and the 9-items display, when overall task complexity increases.

Figure D.12 presents the proportion of misses on the VST as a function of session, age and set size for (a) non-arcsine transformed and (b) arcsine-transformed values, with fitted power functions for the younger age group. The 3-way interaction shown in Figure D.12 including the factors session, age and set size was significant ($F$[3, 96] = 4.02, $\eta_p^2$ = .11, $p < .05$). A follow-up analysis of this interaction with separate 2-way (session x set size) ANOVAs for each age group, revealed a significant interaction for younger adults ($F$[3, 48] = 6.63, $\eta_p^2$ = .29, $p < .01$), but not for older adults ($p > .55$). A follow-up analysis of the significant 2-way interaction for younger adults with separate 1-way (session) ANOVAs for each set size, revealed a simple main effect for set size 9

($F[3, 48] = 3.42$, $\eta_p^2 = .18$, $p < .05$; fitted power function: $y = 0.01 + 0.05x^{-1.31}$, $R^2 = .99$, RMSE = 0.00) as well as for set size 25 ($F[3, 48] = 12.90$, $\eta_p^2 = .45$, $p < .001$; fitted power function: $y = 0.05 + 0.19x^{-0.74}$, $R^2 = .99$, RMSE = 0.01). The interaction represented in Figure D.12 thus shows that with practice, the proportion of misses does not significantly decrease for older adults. However, for younger adults, practice reduces the proportion of misses. As can be seen graphically and statistically with the fitted power functions, the reduction is stronger for the bigger set size as compared to the smaller set size.

The omnibus ANOVA revealed no other significant interactions (all $p$s > .13).

### D.1.3 Subjective Measures

Subjective measures assessed with the RSME and the NASA-TLX rating scales were analyzed with 3-way mixed-factors ANOVAs taking into account the between-participants factor *age* (younger, older) and the within-participants factor *task type* (single-LCT, single-VST, dual, dual-r) to account for differences in mental demand rating according to task complexity as well as the within-participants factor *session* (1-4) to assess the influence of practice on mental demand ratings.

*RSME.* Figure D.13 presents mean RSME as a function of session, age and task type, with fitted power functions for the older adults in the single-VST, the dual and the dual-r condition. The ANOVA revealed a significant main effect of task type on RSME ($F[3, 93] = 47.28$, $\eta_p^2 = .60$, $p < .001$). Participants rated the single-LCT condition as least demanding (M = 20.38), followed by the single-VST condition (M = 28.26), the dual-condition (M = 35.98) and finally the dual-r condition (M = 41.09). There was furthermore a significant main effect of session ($F[3, 93] = 34.84$, $\eta_p^2 = .53$, $p < .001$), with subjective ratings of mental effort decreasing over sessions ($M_1 = 37.02$; $M_2 = 31.03$; $M_3 = 29.26$; $M_4 = 28.39$). The factor age showed a tendency, but missed to reach significance ($p = .07$) indicating that older adults rated the different tasks as more demanding than younger adults. All main effects were modulated by a significant 3-way interaction including the factors task type, session and age ($F[9, 279] = 3.27$, $\eta_p^2 = .10$, $p < .01$). To further analyze this interaction separate 2-way (session x task type) ANOVAs

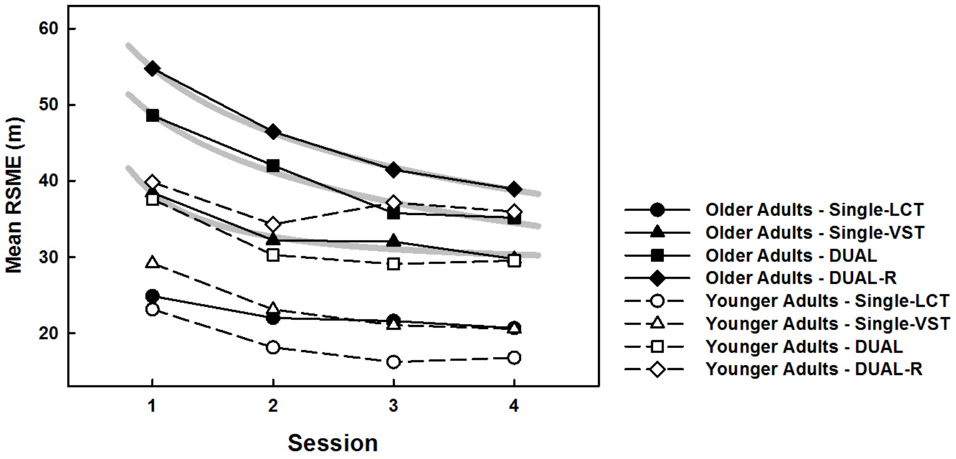


Figure D.13: Mean Rating Scale of Mental Effort (RSME) as a function of session (1-4), age (younger, older) and task type (single-LCT, single-VST, dual, dual-r), with fitted power functions for the older adults in the single-VST, the dual and the dual-r condition.

were computed for younger and older adults respectively. The ANOVAs revealed a significant interaction between those factors for older adults ($F$[9, 135] = 2.83, $\eta_p^2$ = .16, $p$ < .05), but not for younger adults ($p$ > .17). Another post-hoc analysis with separate 1-way (session) ANOVAs for each task type, revealed a simple main effect of session on RSME for the single-VST condition ($F$[3, 48] = 11.07, $\eta_p^2$ = .41, $p$ < .001; fitted power function: $y = 28.73 + 9.66x^{-1.31}$, $R^2$ = .97, RMSE = 1.21), for the dual condition ($F$[3, 48] = 10.68, $\eta_p^2$ = .40, $p$ < .01; fitted power function: $y = -13.77 + 62.49x^{-0.19}$, $R^2$ = .97, RMSE = 1.78) as well as for the dual-r condition ($F$[3, 48] = 12.72, $\eta_p^2$ = .44, $p$ < .001; fitted power function: $y = -15.98 + 70.78x^{-0.19}$, $R^2$ = .99, RMSE = 0.39). The single-LCT condition showed a tendency, but missed to reach significance ($p$ = .08). Looking at Figure D.13 it can thus be said that practice did not change the subjective ratings of mental effort for younger adults, but that for older adults, indeed practice had an effect on subjective rating, especially when performing the single-VST, the dual and dual-r conditions.

The omnibus ANOVA revealed no other significant main effects or interactions ($p$ >

.13).

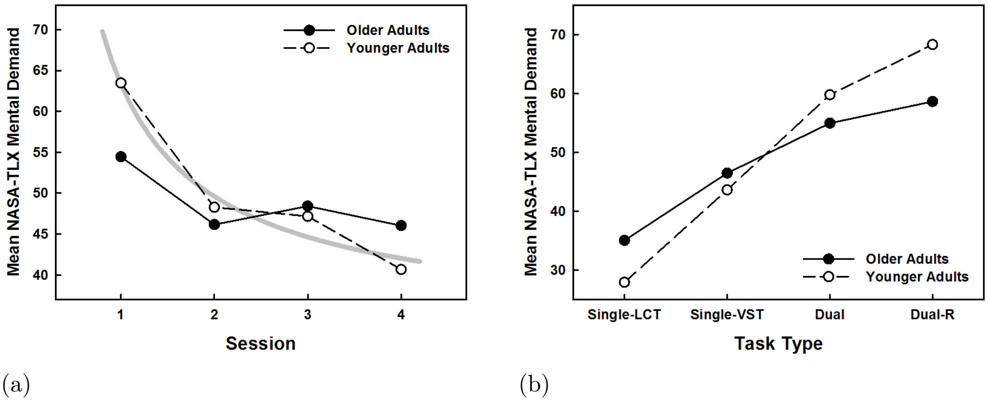
Table D.1: Mean subjective ratings and significance of the main effect of task type on each of the NASA-TLX subscales.

| Category | Single-LCT | Single-VST | DUAL | DUAL-R | Sig. |
|---|---|---|---|---|---|
| Mental Demand | 31.47 | 45.03 | 57.36 | 63.45 | $< .001$ |
| Physical Demand | 25.87 | 26.09 | 41.92 | 48.00 | $< .001$ |
| Temporal Demand | 35.27 | 52.80 | 67.20 | 65.23 | $< .001$ |
| Effort | 95.05 | 97.32 | 86.59 | 87.29 | $< .001$ |
| Performance | 41.29 | 59.98 | 71.33 | 74.36 | $< .001$ |
| Frustration | 20.36 | 29.00 | 33.11 | 34.59 | $< .001$ |

Table D.2: Mean subjective ratings and significance of the main effect of session on each of the NASA-TLX subscales.

| Session | 1 | 2 | 3 | 4 | Sig. |
|---|---|---|---|---|---|
| Mental Demand | 58.97 | 47.21 | 47.79 | 43.34 | $< .001$ |
| Physical Demand | 39.23 | 35.47 | 35.40 | 31.79 | $< .01$ |
| Temporal Demand | 67.55 | 55.14 | 54.16 | 46.67 | $< .001$ |
| Effort | 89.02 | 91.25 | 95.76 | 93.22 | $< .05$ |
| Performance | 73.13 | 65.29 | 57.70 | 53.76 | $< .001$ |
| Frustration | 33.75 | 28.33 | 25.70 | 29.27 | $< .01$ |

*NASA-TLX Mental Demand.* Figure D.14a presents mean NASA-TLX scores for mental demand as a function of session and age, with fitted power function for the younger adults, and Figure D.14b presents mean NASA-TLX scores for mental demand as a function of task type and age. There was a significant main effect of task type on the subjective rating of mental demand ($F[3, 90] = 65.39$, $\eta_p^2 = .69$, $p < .001$). As can be seen in Table D.1 subjective ratings of mental demand were highest in the dual-r condition, followed by the dual-condition, the single-VST condition and finally the single-LCT condition. There was furthermore a significant main effect of session on mental demand



(a) (b)

Figure D.14: Mean NASA-TLX scores for Mental Demand as a function of (a) session (1-4) and age (younger, older), with fitted power function for the younger adults, and (b) task type (single-LCT, single-VST, dual, dual-r) and age (younger, older).

($F[3, 90] = 14.90$, $\eta_p^2 = .33$, $p < .001$): As can be seen in Table D.2 over sessions, ratings of mental demand decreased significantly. The interaction between task type and age was significant ($F[3, 90] = 4.63$, $\eta_p^2 = .13$, $p < .01$). A follow-up with two separate 1-way (task type) ANOVAs for each age group revealed a simple main effect of task type for both younger ($F[3, 45] = 47.58$, $\eta_p^2 = .76$, $p < .001$) as well as older adults ($F[3, 45] = 19.79$, $\eta_p^2 = .57$, $p < .001$). The interaction is explained by the fact that subjective ratings of mental demand increased stronger for younger adults with increasing overall task difficulty as compared to older adults. This might be an indication that older adults over-estimate their own capacities. The interaction between session and age was significant as well ($F[3, 90] = 3.05$, $\eta_p^2 = .09$, $p < .05$). Follow-up analysis with separate 1-way (session) ANOVAs for each age group revealed a simple main effect of session for younger adults ($F[3, 45] = 21.39$, $\eta_p^2 = .59$, $p < .001$; fitted power function: $y = 32.75 + 30.57x$ $^{-0.86}$, $R^2 = .96$, RMSE $= 3.17$) but not for older adults ($p > .14$). The interaction shown in Figure D.14a can thus be explained by the fact that with practice, subjective rating of mental demand decreases significantly for younger adults, but not for older adults. This might again be an indication that older adults subjectively misjudge their own capacities.

The omnibus ANOVA revealed no other significant main effects or interactions (all $p$s $>$ .32).

*NASA-TLX Physical Demand.* There was a significant main effect of task type on the subjective rating of physical demand ($F[3,\ 90] = 34.06$, $\eta_p^2 = .53$, $p < .001$). As can be seen in Table D.1 subjective ratings of physical demand were highest in the dual-r condition, followed by the dual-condition, the single-LCT condition and finally the single-LCT condition. There was furthermore a significant main effect of session on physical demand ($F[3,\ 90] = 5.00$, $\eta_p^2 = .14$, $p < .05$): As can be seen in Table D.2 with practice, ratings of mental demand decreased. The factor age showed a tendency, but did not reach significance ($p = .07$). This might however be an indication that the different tasks were more physically demanding than for younger adults. No other significant interactions could be found (all $p$s $>$ .25).

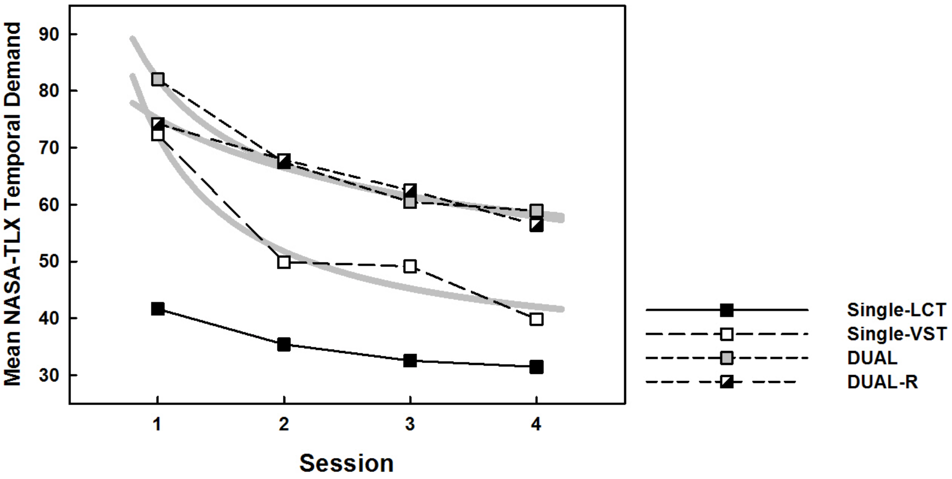


Figure D.15: Mean NASA-TLX scores for Temporal Demand as a function of (a) session (1-4) and task type (single-LCT, single-VST, dual, dual-r), with fitted power functions for the single-VST, dual and dual-r conditions.

*NASA-TLX Temporal Demand.* Figure D.15 presents mean NASA-TLX scores for temporal demand as a function of session and task type, with fitted power functions for the single-VST, dual and dual-r conditions. There was a significant main effect of task type on the subjective rating of temporal demand ($F$[3, 90] = 60.20, $\eta_p^2$ = .67, $p < .001$). As can be seen in Table D.1 subjective ratings of temporal demand were highest in the dual-r condition, followed by the dual-condition, the single-LCT condition and finally the single-LCT condition. There was furthermore a significant main effect of session on mental demand ($F$[3, 90] = 17.98, $\eta_p^2$ = .38, $p < .001$): As can be seen in Table D.2 with practice sessions, ratings of temporal demand decreased. The factor age showed a tendency, but missed to reach significance ($p = .08$). It might indicate however, that older adults felt like they were more under temporal pressure than younger adults. The factor task type interacted significantly with the factor session ($F$[9, 270] = 2.91, $\eta_p^2$ = .09, $p < .05$). A follow-up analysis of this interaction with 4 separate (session) ANOVAs for each task type revealed no simple main effect for single-LCT ($p > .12$), a simple main effect for single-VST ($F$[3, 99] = 14.47, $\eta_p^2$ = .31, $p < .001$; fitted power function: $y = 33.11 + 39.01x^{-1.06}$, $R^2 = .96$, RMSE = 4.86), a simple main effect for the dual condition ($F$[3, 99] = 12.90, $\eta_p^2$ = .28, $p < .001$; fitted power function: $y = 47.63 + 34.46x^{-0.84}$, $R^2 = .99$, RMSE = 1.16) and a simple main effect for the dual-r condition ($F$[3, 99] = 6.52, $\eta_p^2$ = .17, $p < .001$; fitted power function: $y = -2516.05 + 2591.13x^{-0.01}$, $R^2 = .97$, RMSE = 2.45). As can be seen in Figure D.15 the interaction can be explained by the fact that with practice subjective ratings of temporal demand decrease for single-VST, dual- and dual-r conditions. The decrease is however strongest for the single-VST condition and less strong for both dual conditions.

The omnibus ANOVA revealed no other significant interactions (all $p$s > .24).

*NASA-TLX Effort.* There was a significant main effect of task type on the subjective rating of effort ($F$[3, 90] = 7.19, $\eta_p^2$ = .19, $p < .001$). As can be seen in Table D.1 subjective ratings of effort were highest in the dual-r condition, followed by the dual-condition, the single-LCT condition and finally the single-LCT condition. There was furthermore a significant main effect of session on subjective rating of effort ($F$[3, 90] = 3.61, $\eta_p^2$ = .11, $p < .05$): As can be seen in Table D.2 with practice, subjective ratings of effort decreased. The factor age showed a significant main effect on the subjective measure of effort as well ($F$[1, 30] = 4.95, $\eta_p^2$ = .14, $p < .05$): Subjective ratings of effort were lower for older adults (M = 84.17) as compared to younger adults (M = 98.96). This might be an indication that older adults overrate their own capacities. No other significant interactions could be found (all $ps > .16$).

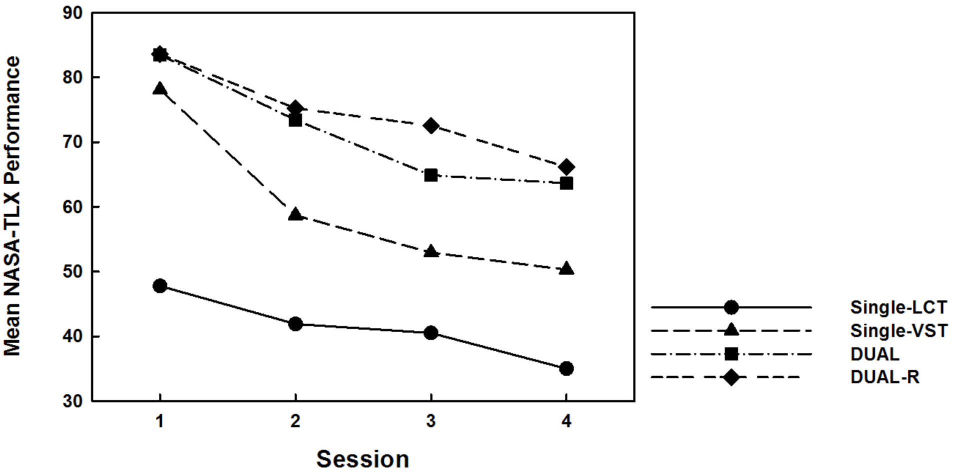


Figure D.16: Mean NASA-TLX scores for Performance as a function of session (1-4) and task type (single-LCT, single-VST, dual, dual-r).

*NASA-TLX Performance.* Figure D.16 presents mean NASA-TLX scores for performance. There was a significant main effect of task type on the subjective rating of performance ($F[3,\ 90] = 53.20$, $\eta_p^2 = .64$, $p < .001$). As can be seen in Table D.1 subjective ratings of performance were highest in the dual-r condition, followed by the dual-condition, the single-LCT condition and finally the single-LCT condition. There was furthermore a significant main effect of session on subjective rating of effort ($F[3,\ 90] = 21.60$, $\eta_p^2 = .42$, $p < .001$): As can be seen in Table D.2 with practice, subjective ratings of performance decreased. The interaction between session and task type showed a strong tendency, but just missed to reach significance ($p = .05$). As can be seen in Figure D.16 however, subjective ratings of performance on some task types seem to decrease stronger than on other tasks. No other significant interactions were found (all $ps > .48$).

*NASA-TLX Frustration.* There was a significant main effect of task type on the subjective rating of frustration ($F[3,\ 90] = 10.15$, $\eta_p^2 = .25$, $p < .001$). As can be seen in Table D.1 the dual-r condition was subjectively rated as most frustrating, followed by the dual-condition, the single-LCT condition and finally the single-LCT condition. There was furthermore a significant main effect of session on the subjective rating of frustration ($F[3,\ 90] = 5.16$, $\eta_p^2 = .15$, $p < .01$): As can be seen in Table D.2 with practice, subjective ratings of frustration decreased. No other significant main effects or interactions could be found (all $p$s $> .24$).

## D.2 Analyzing Dual-Task Data in Detail

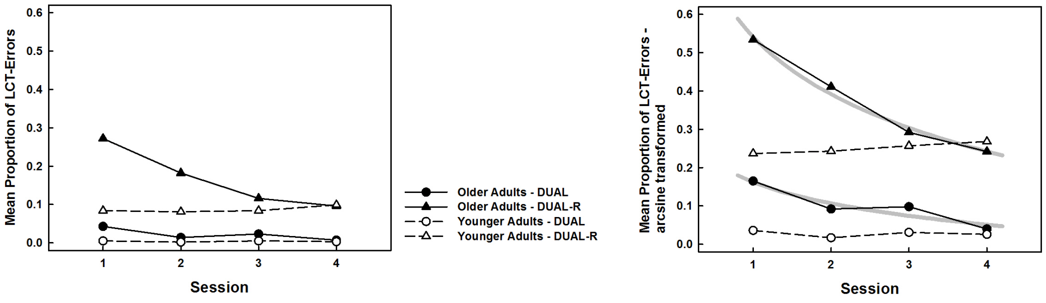
### D.2.1 Driving Measures

In this section we have a closer look at dual-task driving data. Separate 4-way mixed-factors ANOVAs were conducted on the driving measures LCT-Errors, LCT-Misses, RT-LC, MT-LC, MDEV, SDDEV, SDSW and IRI. The ANOVA included the within-participants factor *set size* (9, 25) taking into account the effect of secondary task difficulty, the within-participants factor *session* (1-4) to analyze the effect of repetition over sessions, the within-participants factor *dual-task type* (dual, dual-r) to account for differences in dual-task setting and finally the between-participants factor *age* (younger, older) to account for age effects. All measures were based on LC segments only. To avoid partial redundancy with the data presented in the section above, for the driving measures RT-LC, MT-LC, MDEV, SDDEV and SDSW only relevant new main effects and interactions including the factors dual-task type and set size are reported here.

*LCT-Errors.* Figure D.17 presents the mean proportion of LCT-Errors as a function of session, dual-task type and age, with fitted power functions for the older age group. Figure D.18 presents the mean proportion of LCT-Errors as a function of set size and age. The ANOVA revealed a significant main effect of dual-task type ($F[1,\ 32] = 173.89$, $\eta_p^2 = .85$, $p < .001$). LCT-errors were significantly higher in the dual-r condition (M = .13) as compared to the dual condition (M = .01). There was furthermore a significant main effect of session on LCT-Errors ($F[3,\ 96] = 12.20$, $\eta_p^2 = .28$, $p < .001$): With practice LCT-Errors decreased. There was a significant main effect of set size as well ($F[1,\ 32] = 9.59$, $\eta_p^2 = .23$, $p < .01$): The proportion of LCT-Errors was significantly higher with set size 25 (M = .08) as compared to set size 9 (M = .06). There was finally a main effect of age on the proportion of LCT-Errors ($F$[1, 32] = 11.54, $\eta_p^2$ = .27, $p$ < .01): Older adults made significantly more LCT-Errors (M = .09) in comparison to younger adults (M = .05).

The factor session interacted significantly with the factor age ($F$[3, 96] = 15.54, $\eta_p^2$ = .33, $p$ < .001), but this 2-way interaction was modulated by a significant 3-way interaction including the factors session, age and dual-task type ($F$[3, 96] = 3.51, $\eta_p^2$ = .10, $p$ < .05), represented in Figure D.17. A follow-up analysis of this interaction with separate 2-way (session x age) ANOVAs for the dual- and dual-r condition respectively, revealed a significant interaction between session and age for the dual condition ($F$[3, 96] = 3.85, $\eta_p^2$ = .11, $p$ < .05) as well as for the dual-r condition ($F$[3, 96] = 8.59, $\eta_p^2$ = .21, $p$ < .01). Follow-ups of each 2-way (session x age) ANOVA with separate 1-way (session) ANOVAs for younger and older adults respectively and for each dual-task type condition, revealed for the dual-condition a simple main effect of session for older adults ($F$[3, 48] = 4.60, $\eta_p^2$ = .22, $p$ < .05; fitted power function: $y$ = 1.31 - 1.15$x^{-0.0667}$, $R^2$ = .88, RMSE = 0.02), but not for younger adults ($p$ > .46). In the dual-r condition, there was a simple main effect of session as well for older adults ($F$[3, 48] = 9.73, $\eta_p^2$ = .38, $p$ < .01; fitted power function: $y$ = 1.19 - 0.65$x^{-0.2748}$, $R^2$ = .99, RMSE = 0.03), but not for younger adults ($p$ > .66). The interaction shown in Figure D.17 can thus be explained as follows: The proportion of LCT-Errors for younger adults does not change with practice whether in the dual- or the dual-r condition. However, for older adults, practice has a positive effect and the proportion of LCT-Errors decreases over time for both the dual- as well as the dual-r condition. The decrease in LCT-Errors is stronger for the dual-r condition as compared to the dual condition.

There was furthermore a significant interaction between the factors set size and age ($F$[1, 32] = 6.62, $\eta_p^2$ = .17, $p$ < .05). A follow-up analysis with separate 1-way (set size) ANOVAs for each age group, revealed a simple main effect of set size for older adults ($F$[1, 16] = 18.49, $\eta_p^2$ = .54, $p$ < .01), but not for younger adults ($p$ > .62). The interaction shown in Figure D.18 can thus be explained by the fact that with increasing set size, the proportion of LCT-Errors increases significantly for older adults, but not for younger adults.



(a) (b)

Figure D.17: Proportion of LCT-Errors (a) session (1-4), age (younger, older) and dual-task type (dual, dual-r) for (a) non-arcsine transformed and (b) arcsine-transformed values, with fitted power functions for the older age group.

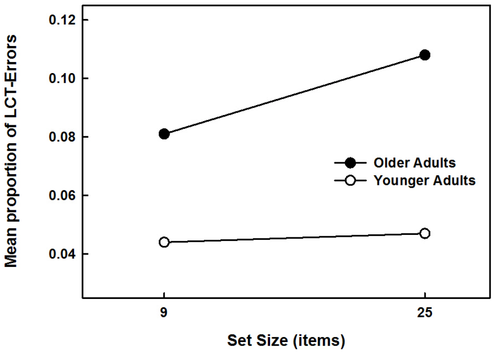
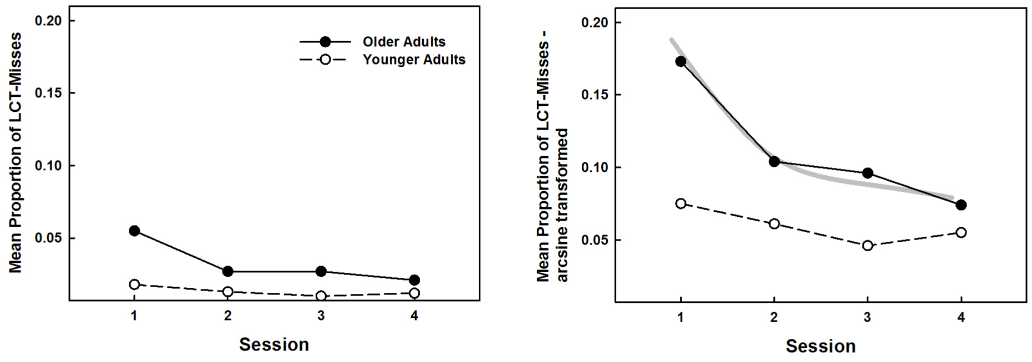


Figure D.18: Proportion of LCT-Errors as a function of set size (9, 25) and age (younger, older).

The omnibus ANOVA revealed no other significant interactions (all $p$s $>$ .12).

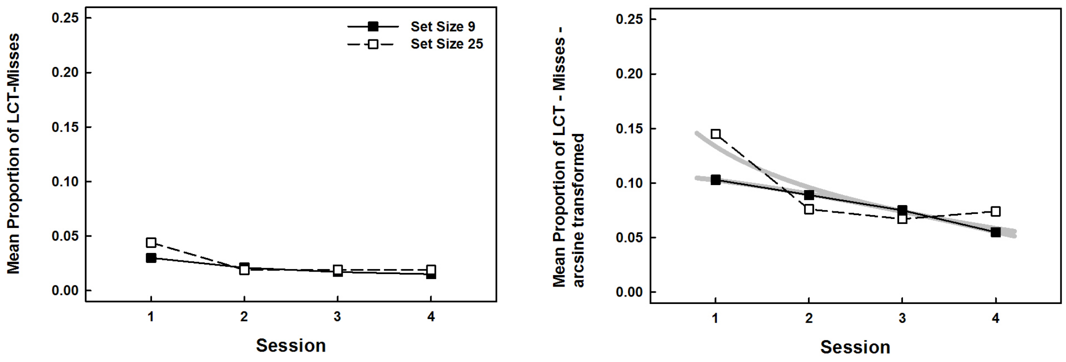
*LCT-Misses.* Figure D.19 presents the proportion of LCT-Misses as a function of session and age for (a) non-arcsine transformed values and (b) arcsine-transformed values, with fitted power function for the older adults. Figure D.20 presents the proportion of LCT-Misses as a function of session and set size for (a) non-arcsine transformed and (b) arcsine-transformed values with fitted power functions. Figure D.21 presents the proportion of LCT-Misses as a function of session and dual-task type for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions. The ANOVA revealed a significant main effect of dual-task type on LCT-Misses ($F$[1, 32] = 99.97, $\eta_p^2$ = .76, $p <$ .001). Significantly more LCT-Misses were produced in the dual-r condition (M = .04) as compared to the dual condition (M = .01). There was a significant main effect of session on LCT-Misses as well ($F$[3, 96] = 7.86, $\eta_p^2$ = .20, $p <$ .001) indicating that with practice the number of LCT-misses decreased. There was finally a significant main effect of age on LCT-Misses ($F$[1, 32] = 8.96, $\eta_p^2$ = .22, $p <$ .01): Older adults missed more lane changes (M = .03) than younger adults (M = .01). There was furthermore a significant interaction between session and age ($F$[3, 96] = 2.93, $\eta_p^2$ = .08, $p <$ .05). A follow-up analysis with two separate 1-way (session) ANOVAs for younger and older adults respectively, revealed a simple main effect of session for older adults ($F$[3, 48] = 6.29, $\eta_p^2$ = .28, $p <$ .01; fitted power function: $y = 0.05 + 0.12x^{-1.14}$, $R^2$ = .98, RMSE = 0.01), but not for younger adults ($p >$ .15). As can be seen in Figure D.19, the interaction comes from the fact that with practice, the proportion of LCT-Misses decreases significantly for older adults, but not for younger adults.

The 2-way interaction between the factors session and set size shown in Figure D.20 was significant as well ($F$[3, 96] = 3.08, $\eta_p^2$ = .09, $p <$ .05). A post-hoc analysis with separate 1-way (session) ANOVAs for set size 9 and set size 25 respectively revealed a simple main effect for set size 9 ($F$[3, 99] = 3.58, $\eta_p^2$ = .10, $p <$ .05; fitted power function: $y = 0.11 - 0.01x^{-1.45}$, $R^2$ = .99, RMSE = 0.00) as well as set size 25 ($F$[3, 99] = 6.48, $\eta_p^2$ = .16, $p <$ .01; fitted power function: $y = 319.62 - 319.48x^{-0.0002}$, $R^2$ = .79, RMSE = 0.93). The interaction shown in Figure D.20 can be explained by the fact that with practice, the proportion of LCT-Misses decreases stronger for the set size with 25 items as compared to the smaller set size with 9 items.



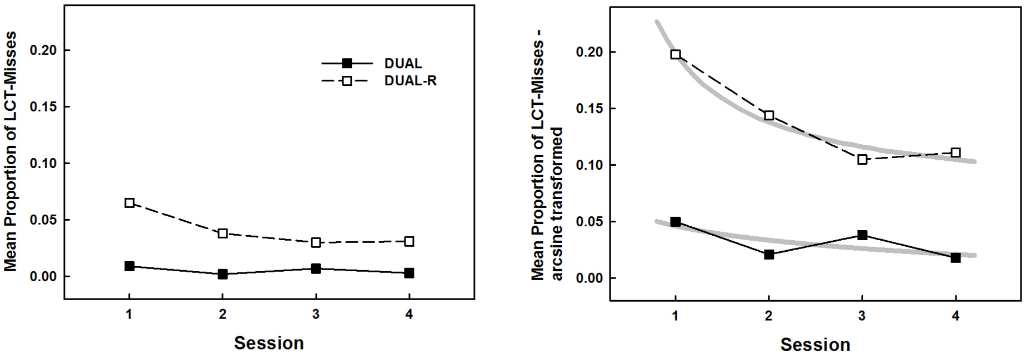
(a) (b)

Figure D.19: Proportion of LCT-Misses as a function of session (1-4) and age (younger, older) for (a) non-arcsine transformed values and (b) arcsine-transformed values, with fitted power function for the older adults.



(a) (b)

Figure D.20: Proportion of LCT-Misses as a function of session (1-4) and set size (9, 25) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions.



(a) (b)

Figure D.21: Proportion of LCT-Misses as a function of session (1-4) and dual-task type (dual, dual-r) for (a) non-arcsine transformed values and (b) arcsine-transformed values with fitted power functions.

Finally, the ANOVA revealed a significant interaction between the factors session and dual-task type ($F[3,\ 96] = 3.32$, $\eta_p^2 = .09$, $p < .05$). A post-hoc analysis with separate 1-way (session) ANOVAs for dual and dual-r, revealed a simple main effect of session for dual ($F[3,\ 99] = 3.21$, $\eta_p^2 = .09$, $p < .05$; fitted power function: $y = 65.32 - 65.27x^{-0.0003}$, $R^2 = .53$, RMSE $= 5.26$) as well as dual-r ($F[3,\ 99] = 5.93$, $\eta_p^2 = .15$, $p < .01$; fitted power function: $y = 0.06 + 0.13x^{-0.8700}$, $R^2 = .96$, RMSE $= 0.01$). The interaction shown in Figure D.21 can be explained by the fact that with practice, the proportion of LCT-Misses decreases stronger in the dual-r condition as compared to the dual-condition.

The omnibus ANOVA revealed no other significant main effects or interactions (all $p$s $> .18$).

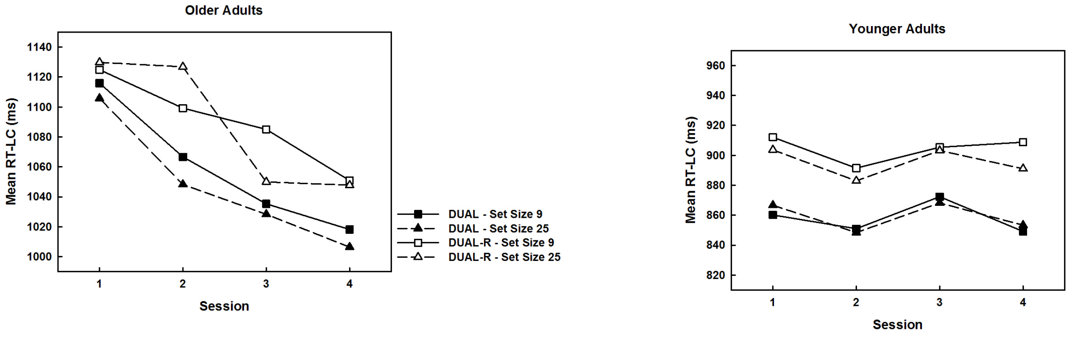
*RT-LC.* Figure D.22a presents mean RT-LC as a function of session, dual-task type and set size for older adults and Figure D.22b presents mean RT-LC as a function of session, dual-task type and set size for younger adults. The ANOVA showed a significant main effect of dual-task type on RT-LC ($F[1,\ 32] = 28.127$, $\eta_p^2 = .47$, $p < .001$). RT-LC was significantly higher in the dual condition (M = 955 ms) as compared to the dual-r condition (M = 995 ms). The factor set size showed a tendency but missed to reach significance ($p$ = .09).

The 3-way interaction between dual-task type, set size and age ($p$ = .09) showed a tendency, but missed to reach significance. The 4-way interaction shown in Figure D.22 including the factors session, dual-task type, set size and age however reached significance ($F[3, 96] = 3.26$, $\eta_p^2 = .09$, $p < .05$). To analyze this interaction, two separate 3-way ANOVAs including the factors session, dual-task type and set size were conducted for each age group. For the younger adults the 3-way interaction session x dual-task type x set size was not significant ($p > .40$). For the older adults, this same interaction was significant ($F[3, 48] = 3.43$, $\eta_p^2 = .18$, $p < .05$). A further post-hoc analysis of this significant 3-way interaction with separate 2-way (session x setsize) ANOVAs for the dual and the dual-r condition respectively, revealed no significant interaction for the dual condition ($p > .62$) nor for the dual-r condition ($p > .21$). The interaction shown in Figure D.22a can thus be explained as follows: RT-LC values for older adults decrease over sessions. Although the 2-way interactions (session x set size) were not significant, patterns differ for the two task types (dual, dual-r), as revealed by the significant 3-way interaction including the factors session, dual-task type and set size. As can be seen in Figure D.22a, RT-LC values decrease more or less in parallel for both set sizes in the dual-condition, whereas in the dual-r condition, RT-LC decreases steadily for the 9-item set size, but not for the 25-item set size.

The omnibus ANOVA revealed no other significant main effects or interactions with dual-task type or set size (all $p$s $> .11$).

*MT-LC.* The interaction between the factors dual-task type, set size and age showed a strong tendency, but missed to reach significance ($p$ = .06). No significant main effects or interactions could be found for MT-LC (all $p$s $> .12$).

*MDEV.* Figure D.23 presents mean MDEV as a function of dual-task type, set size and age. The ANOVA showed a significant main effect of dual-task type on MDEV ($F[1, 32] = 31.22$, $\eta_p^2 = .50$, $p < .001$). Mean MDEV values were significantly higher in the dual-task-R condition (M = 1.92 m) as compared to the dual condition (M = 1.83



(a) (b)

Figure D.22: Mean Reaction Time until Lane Change (RT-LC) as a function of session (1-4), dual-task type (dual, dual-r) and set size (9, 25) for (a) older adults and (b) younger adults.

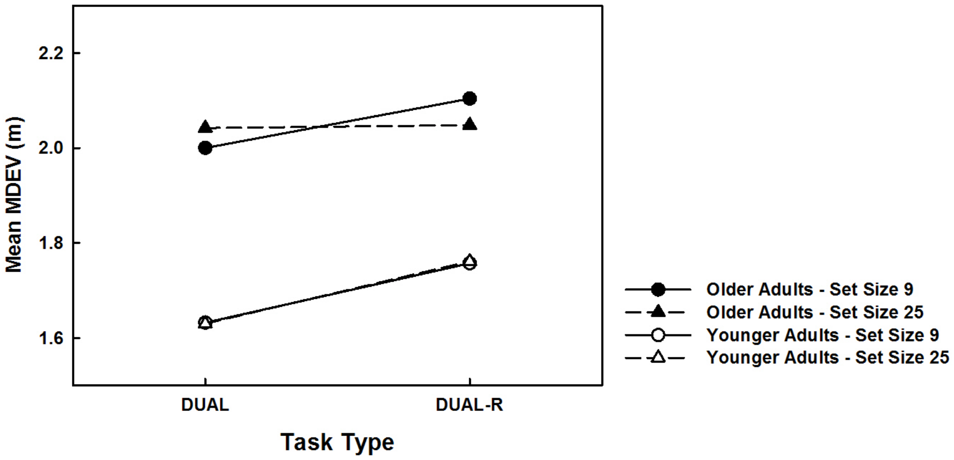


Figure D.23: Mean Lateral Deviation (MDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25).

m). The interaction between dual-task type and age was significant as well ($F[1, 32] = 5.08$, $\eta_p^2 = .14$, $p < .05$). The interaction between dual-task type and set size showed a tendency but missed to reach significance ($p = .06$). Those 2-way interactions were however all modulated by a significant 3-way interaction between the factors dual-task type, set size and age ($F[1, 32] = 4.69$, $\eta_p^2 = .13$, $p < .05$). A post-hoc analysis with separate 2-way (dual-task type x set size) ANOVAs for each age group, revealed that this interaction was significant for older adults ($F[1, 16] = 7.01$, $\eta_p^2 = .31$, $p < .05$), but not for younger adults ($p > 88.$). A follow-up on the 2-way interaction for older adults with two separate 1-way (dual-task type) ANOVAs for set size 9 and 25 respectively, showed that there was a significant simple main effect of task type for the set size of 9 items ($F[1, 16] = 15.28$, $\eta_p^2 = .49$, $p < .01$), but not for the set size of 25 items ($p > 89.$). This means that for older adults, when the 9-item visual search display was presented, MDEV values increased significantly in dual-r conditions as compared to dual conditions.

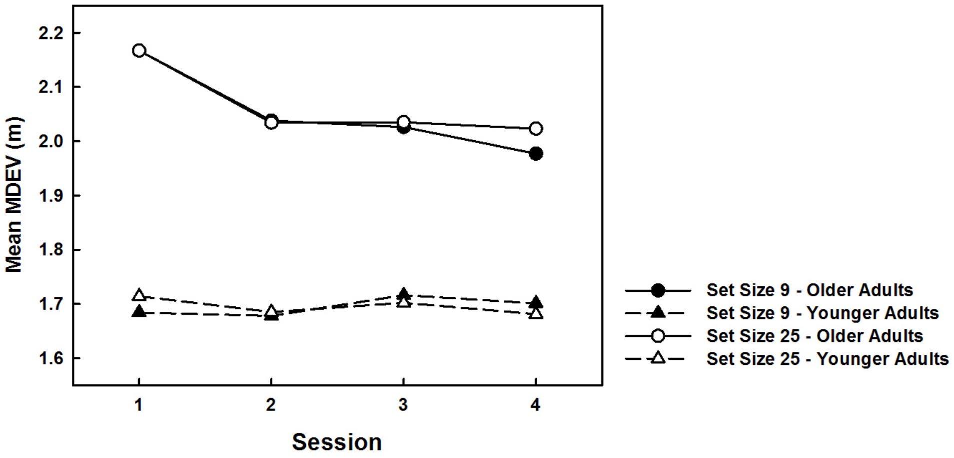


Figure D.24: Mean Lateral Deviation (MDEV) as a function of session (1-4), age (younger, older) and set size (9, 25).

Figure D.24 presents mean MDEV as a function of session, age and set size. The 3-way interaction between those factors was significant ($F[3, 96] = 2.92$, $\eta_p^2 = .08$, $p < .05$). A follow-up analysis with separate 2-way (session x set size) ANOVAs for each age group, revealed a strong tendency for the older adults ($p = .06$) and no interaction for the younger adults ($p > .60$). As can be seen in Figure D.24 MDEV values do not change with practice for younger adults, but for older adults, MDEV performance does seem to change with the smaller set size.

The omnibus ANOVA revealed no other significant main effects or interactions with dual-task type or set size (all $p$s $> .29$).

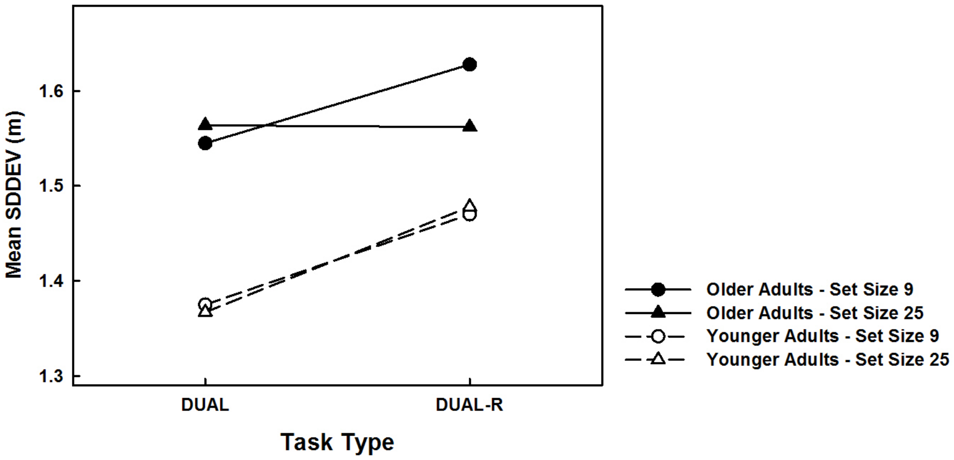


Figure D.25: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of dual-task type (dual, dual-r), age (younger, older) and set size (9, 25).

*SDDEV*. Figure D.25 presents mean SDDEV as a function of dual-task type, age and set size. The ANOVA showed a significant main effect of dual-task type on SDDEV ($F[1,\ 32] = 31.11$, $\eta_p^2 = .49$, $p < .001$). Mean SDDEV values were significantly higher in the dual-r condition (M = 1.54 m) as compared to the dual condition (M = 1.46 m). The factor dual-task type interacted significantly with the factor age as well ($F[1,\ 32] = 5.82$, $\eta_p^2 = .15$, $p < .05$), but this interaction was modulated by a significant 3-way interaction between the factors dual-task type, set size and age ($F[1,\ 32] = 5.51$, $\eta_p^2 = .15$, $p < .05$). To follow-up on this interaction separate 2-way (dual-task type x set size) ANOVAs were conducted for younger and older adults respectively. The interaction was significant for older adults ($F[1,\ 16] = 7.77$, $\eta_p^2 = .33$, $p < .05$), but not for younger adults ($p > .60$). A further analysis of the significant interaction for older adults with separate 1-way (dual-task type) ANOVAs for each set size, revealed a simple main effect of dual-task type for set size 9 ($F[1, 16] = 16.26$, $\eta_p^2 = .50$, $p < .01$), but not for set size 25 ($p > .96$). As can be seen in Figure D.25 older adults suffer from an increase in dual-task complexity as can be seen by an increase in SDDEV values, but only when the visual-search display with 9 items is presented. Younger adults' SDDEV performance is not affected by dual-task type complexity.

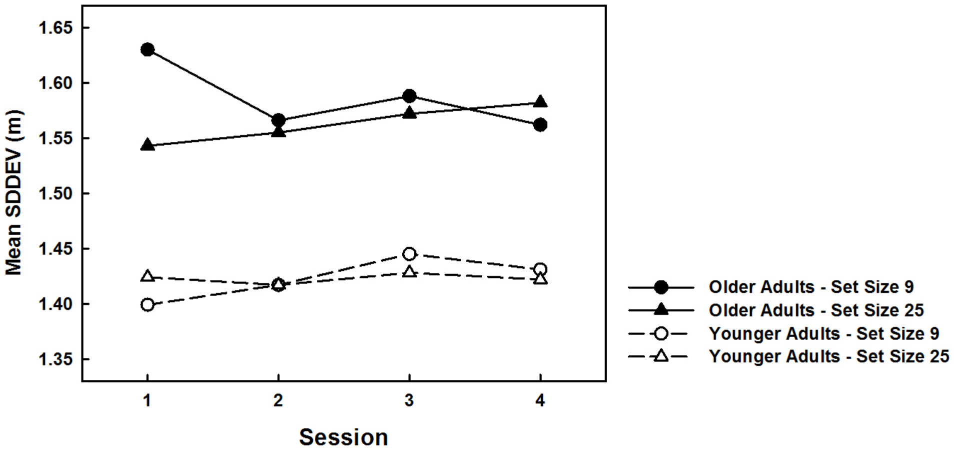
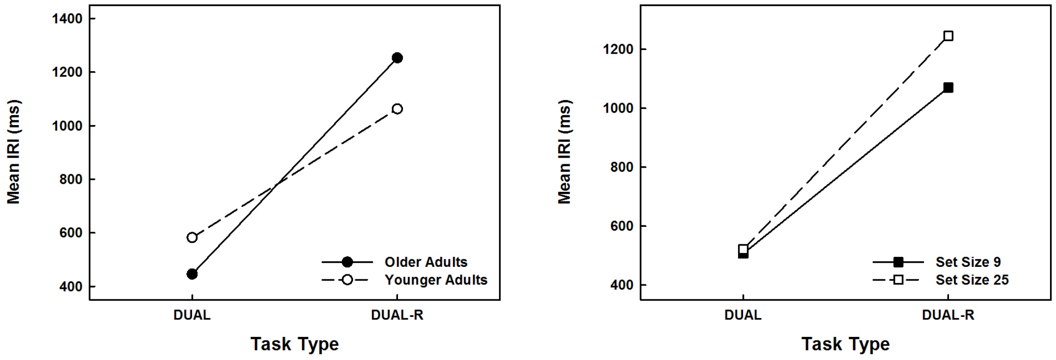


Figure D.26: Mean Standard Deviation of Lateral Deviation (SDDEV) as a function of session (1-4), age (younger, older) and set size (9, 25).

Figure D.26 presents mean SDDEV as a function of session, age and set size. The 3-way interaction between those factors was significant ($F[3, 96] = 3.16$, $\eta_p^2 = .09$, $p < .05$). A follow-up on this interaction with separate 2-way (session x set size) ANOVAs for each age group respectively, revealed a significant interaction of those factors for older adults ($F[3, 48] = 3.356$, $\eta_p^2 = .17$, $p < .05$), but not for younger adults ($p > .59$). A further analysis of this significant interaction with separate 1-way (session) ANOVAs for set size 9 and 25 respectively, revealed no significant effect for set size 9 ($p > .13$) nor for set size 25 ($p > .62$). Looking at Figure D.26 it can thus be said that practice had an effect on SDDEV values for older adults, but not for younger adults.

The omnibus ANOVA revealed no other significant main effects or interactions with



(a) (b)

Figure D.27: Mean Inter-Response Interval (IRI) as a function of (a) dual-task type (dual, dual-r) and age (younger, older) and (b) dual-task type (dual, dual-r) and set size (9, 25).

dual-task type or set size (all *p*s $>$ .12).

*SDSW.* The main effect of dual-task type showed a tendency, but just missed to reach significance ($p$ = .05). Although not statistically significant, the analysis shows that steering-wheel variability was higher in the dual condition (M = 12.58°) as compared to the dual-r condition (M = 12.22°). No other significant main effects or interactions with the factors dual-task type or set size could be found (all *p*s $>$ .15).

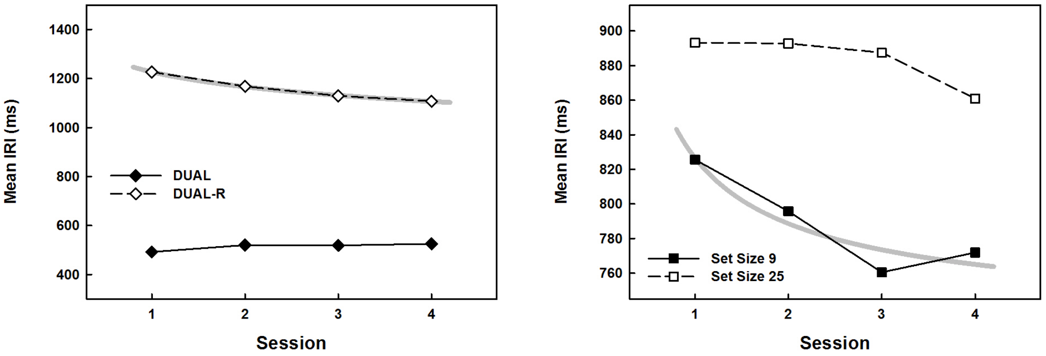
*IRI.* Figure D.27a presents mean IRI as a function of dual-task type and age and Figure D.27b presents mean IRI as a function of dual-task type and set size. The ANOVA showed a significant main effect of dual-task type on IRI ($F$[1, 32] = 31.11, $\eta_p^2$ = .49, $p$ $<$ .001). Mean IRI values were significantly higher in the dual-task-R condition (M = 1158 ms) as compared to the dual condition (M = 514 ms). There was furthermore a significant main effect of set size on IRI ($F$[1, 32] = 104.70, $\eta_p^2$ = .77, $p < .001$): IRI was significantly higher for the 25-items display (M = 884 ms) as compared to the 9-items display (M = 788 ms).

The 2-way interaction between dual-task type and age shown in Figure D.27a was significant ($F[1,\ 32] = 78.86$, $\eta_p^2 = .71$, $p < .001$). Follow-up tests with separate 1-way (dual-task type) ANOVAs for older and younger adults respectively, revealed a simple main effect of dual-task type for older adults ($F[1,\ 16] = 613.51$, $\eta_p^2 = .98$, $p < .001$) as well as for younger adults ($F[1,\ 16] = 704.94$, $\eta_p^2 = .98$, $p < .001$). Independent samples $t$-tests comparing IRI values between age groups for the dual-, as well as the dual-r condition, revealed that for both conditions, differences between age groups were significant (dual: $t[32] = 2.59$, $p < .01$; dual-r: $t[32] = 2.98$, $p < .01$). The interaction comes from the fact that with increasing task difficulty IRI increases stronger for older adults as compared to younger adults.

The 2-way interaction between dual-task type and set size shown in Figure D.27b was significant as well ($F[1,\ 32] = 36.64$, $\eta_p^2 = .53$, $p < .001$). A follow-up analysis with separate 1-way (dual-task type) ANOVAs for set size 9 and 25 respectively, revealed a significant main effect for the visual search display containing 9 items ($F[1,\ 33] = 313.33$, $\eta_p^2 = .91$, $p < .001$) as well as the visual search display containing 25 items ($F[1,\ 33] = 322.11$, $\eta_p^2 = .91$, $p < .001$). The interaction shown in Figure D.27b comes from the fact that with an increase in dual-task complexity, IRI increases stronger with the 25-items set size as compared to the 9-items set size.

Figure D.28a presents mean IRI as a function of session and dual-task type, with fitted power function for dual-r. The factor session interacted significantly with the factor dual-task type ($F[3,\ 96] = 16.11$, $\eta_p^2 = .34$, $p < .001$). Follow-up tests with separate 1-way (session) ANOVAs for dual and dual-r respectively, revealed a simple main effect of session for the dual-r condition ($F[3,\ 99] = 8.43$, $\eta_p^2 = .20$, $p < .001$; fitted power function: $y = \text{-}3538.95 + 4766.04x^{\text{-}0.02}$, $R^2 = .99$, RMSE $= 2.99$), but not for the dual condition ($p > .36$). The interaction shown in Figure D.28a thus comes from the fact that with practice, IRI values decrease for dual-r conditions, but not for dual conditions. This finding can be explained by the fact that for the dual condition, participants were maybe already at a performance maximum with little room for improvement.

Figure D.28b presents mean IRI as a function of session and set size with fitted power function for set size 9. The factor session interacted significantly with the factor set size ($F[3,\ 96] = 3.64$, $\eta_p^2 = .10$, $p < .05$). A follow-up test with separate 1-way ANOVAs



(a)

(b)

Figure D.28: Mean Inter-Response Interval (IRI) as a function of (a) session (1-4) and dual-task type (dual, dual-r), with fitted power function for dual-r and (b) session (1-4) and set size (9, 25), with fitted power function for set size 9.

(session) for set size 9 and 25 respectively, revealed a simple main effect for set size 9 ($F[3, 99] = 3.62$, $\eta_p^2 = .10$, $p < .05$; fitted power function: $y = 726.40 + 100.18x^{-0.69}$, $R^2 = .89$, RMSE = 16.34), but not for set size 25 ($p > .43$). This indicates that practice has a positive effect on dual-task interference when the less complex visual search display is shown. Practice has no effect on dual-task interference with the more difficult visual search display including 25 items.

Figure D.29 presents mean IRI as a function of set size and age. Both factors showed a significant interaction ($F[1, 32] = 24.67$, $\eta_p^2 = .44$, $p < .001$). Follow-up tests with separate 1-way (set size) ANOVAs for older and younger adults respectively, revealed a simple main effect for older ($F[1, 16] = 9.55$, $\eta_p^2 = .37$, $p < .01$) as well as younger adults ($F[1, 16] = 210.47$, $\eta_p^2 = .93$, $p < .001$). The interaction can be explained by the fact that an increase in set size has stronger effect on mean IRI values for younger adults as compared to older adults.

The 3-way interaction represented in Figure D.30 between session, set size and age showed a tendency, but missed to reach significance ($p = .07$). It seems to indicate however that practice has a positive effect on IRI values for younger adults (for both

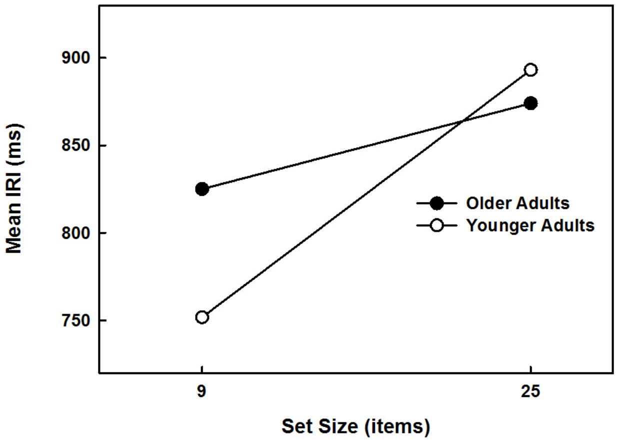


Figure D.29: Mean Inter-Response Interval (IRI) as a function of set size (9, 25) and age (younger, older).

set sizes). For older adults, IRI values remain relatively unchanged for the smaller set size under the effect of practice and even numerically increase a bit for the larger set size.

The omnibus ANOVA revealed no other significant interactions with dual-task type or set size (all *p*s $> .43$).

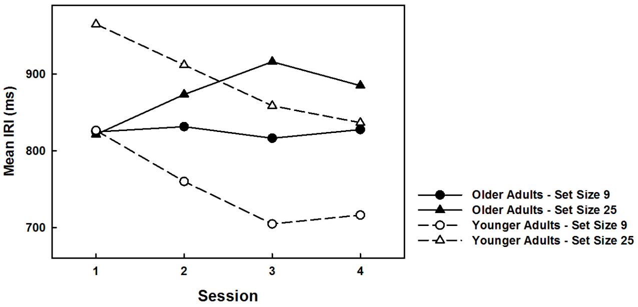


Figure D.30: Mean Inter-Response Interval (IRI) as a function of session (1-4), age (younger, older) and set size (9, 25).