

Driver Behaviour in Level 3 Automated Vehicles: The Effects of Distraction, Fatigue and driver Age

A thesis submitted to
Trinity College Dublin, The University of Dublin,
for the degree of
Doctor of Philosophy

Bilal Alam Khan
School of Psychology
Trinity College Dublin, The University of Dublin

July 2025

DECLARATION

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Summary

This thesis investigated the factors influencing driver takeover performance (TOP) in Level 3 automated vehicles through a series of studies. The investigation began with a foundational systematic literature review (Chapter 1) that systematically examined existing knowledge about older drivers' TOP capabilities, identifying significant gaps in understanding how cognitive ageing processes interact with autonomous driving systems. This review established the theoretical framework for subsequent experimental work.

Building on these insights, Chapter 2 developed and tested a novel conceptual model linking the Takeover Request (TOR) paradigm to core executive functions - specifically Updating, Shifting, and Inhibition capacities. The experimental results revealed important age-related differences: while older adults demonstrated slower reaction times and committed more errors across all cognitive tasks (particularly in inhibition tasks), their performance showed meaningful improvement with repeated exposure. This finding suggested that targeted training interventions could help mitigate age-related challenges.

The investigation then progressed to examining TOR in a high-fidelity driving simulator in Chapter 3, which compared takeover performance under different time budgets (2.5s vs 5s). This study yielded several key insights: longer time budgets significantly improved lane-keeping precision and steering control, yet surprisingly had no measurable impact on drivers' initial response times. Notably, while older drivers reported higher subjective workload ratings, their actual TOP metrics remained comparable to younger drivers across all time budget conditions. This dissociation between subjective experience and objective performance warrants further investigation.

Chapter 4 expanded the scope to consider how non-driving related tasks (NDRTs) influence TOP. The results demonstrated that while response takeover time (RTOT) remained unaffected, engagement in NDRTs – particularly passive monitoring tasks – substantially degraded vehicle control stability (lane keeping).

The final experimental chapter (Chapter 5) investigated the compounded effects of fatigue and NDRT engagement on TOP. The findings revealed a complex relationship: while fatigue significantly impaired vehicle control precision (steering controls), moderate NDRT engagement interestingly helped maintain driver alertness and partially offset fatigue-related performance declines. However, this benefit came with an important caveat – excessive NDRT engagement ultimately degraded lane-keeping performance, highlighting the need for carefully calibrated workload management systems in automated vehicles.

Acknowledgement

First and foremost, I extend my deepest gratitude to God for granting me the strength, wisdom, and opportunities to complete this journey.

I am profoundly grateful to my supervisors, whose mentorship has been instrumental in shaping this thesis. Dr. Sam Cromie, your encouragement and positive feedback gave me the confidence to explore new questions, even when they seemed completely outlandish. Yet, you never dismissed them outright; instead, you encouraged me to pursue them, allowing me to grow as a researcher. Dr. Michael Gormley, your keen insights and constructive critiques provided a much-needed reality check, ensuring that my explorations remained grounded and purposeful, ultimately strengthening this work. Dr. Maria Chiara Leva, your passion for always striving to do more, combined with your kind spirit has always encouraged me – sometimes with just the right amount of insistence – to step beyond my comfort zone. Your efforts in providing invaluable opportunities, resources, and experiences were fundamental to the completion of this PhD.

This thesis was conducted with the financial support of the Research Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224. I would also like to extend my gratitude to Dr. Stephen Carroll for his invaluable support as the programme manager for D-REAL, where I conducted my PhD research. Additionally, I am grateful to Paula Hicks and Rory Carrick for their assistance in facilitating various administrative matters.

I would also like to express my heartfelt appreciation to Anam – a dear friend. Your unwavering support has been my anchor, and without you, this thesis would have been lost. Thank you for standing by me through every challenge, for your patience, and for making this journey as much yours as it is mine.

To my parents, I am forever grateful for your sacrifices, for believing in me, and for standing by me even when I tested your patience. Your unwavering support has been my foundation.

Lastly, to my friends – Farhan, Wasif, Sahil, Mirza Belal, Rahat, Khan Jazib, Muhammad Ali, Waseem Shariff and others who aren't mentioned here – who looked out for me, took care of me in countless ways, offered their encouragement, and reminded me that I was never alone – your kindness created the space in which this work could grow and flourish.

To each and every one of you, thank you!

PUBLICATIONS ARISING FROM THE PRESENT WORK

Khan, Bilal, Coyne Rory, Kielty Paul, Lemley Joseph, Gormley Michael, Leva Maria Chiara, Cromie Sam, “*How Does Age Influence Fatigue Development and Perception in Level 3 Autonomous Driving?*”, (Submitted 29 March 2025, Journal: Transportation Research Part F: Traffic Psychology and Behaviour focuses, IF 3.5). (Under editorial review)

This paper corresponds to chapter 5.

Khan, Bilal, Kielty Paul, Lemley Joseph, Gormley Michael, Leva Maria Chiara, Cromie Sam, “*Age Differences in Takeover Performance Across Varying Time Budgets in Level 3 Autonomous Vehicles*”, (Submitted 27 February 2025, Journal: Human Factors: The Journal of the Human Factors and Ergonomics Society, IF 2.99). (Under review)

This paper corresponds to chapter 3.

Khan, Bilal, Cromie Sam, Leva Maria Chiara, Gormley, Michael, “*An Investigation of Updating, Shifting, and Inhibition Components of Executive Function in Younger and Older Adults – Implications for Takeover Request Functionality in Autonomous Vehicles*”, Available at SSRN: <https://ssrn.com/abstract=4824834> or <http://dx.doi.org/10.2139/ssrn.4824834> (Under review)

This paper corresponds to chapter 2.

Khan B. A., Leva M.C., Cromie S., “*A Systematic Review of Older Drivers in a Level 3 Autonomous Vehicle: A Cognitive Load Perspective*”. In: Longo L., Leva M.C. (eds) Human Mental Workload: Models and Applications. H-WORKLOAD 2021. Communications in Computer and Information Science, vol 1493, 2021.

This paper corresponds to chapter 1.

Abstract

As Level 3 autonomous vehicles (AVs) require drivers to resume control upon a Takeover Request (TOR), understanding the factors influencing takeover performance (TOP) is critical for safety. This thesis examined cognitive, demographic, and situational variables – including age, time budget (TB), Non-Driving Related Tasks (NDRTs), and fatigue – to determine their impact on TOP. This thesis employed a multi-method approach, including a systematic literature review, cognitive experiments, and driving simulations using hi-fidelity simulator, to assess how these variables impact TOP metrics like reaction time, steering behaviour, and lane-keeping performance.

The experimental design used a mixed-factorial approach, with TB, NDRTs, and fatigue as within-subject conditions and age as between-subject factors. Objective driving performance metrics – including response takeover time (RTOT), steering wheel angle (SWA), steering wheel speed (SWS), lateral and longitudinal deceleration, and standard deviation of lane position (SDLnP) – evaluated takeover quality and timeliness. Subjective workload measures (NASA-TLX, Karolinska Sleepiness Scale) assessed cognitive demand and fatigue.

Key findings revealed that while older adults exhibited slower cognitive responses in lab tasks, their real-world TOP remained comparable to that of younger drivers, likely due to compensatory behaviours like driving at lower speed. TB influenced steering precision but not RTOT, suggesting reflexive responses remained stable regardless of warning duration. NDRTs disrupted vehicle control, with passive tasks (e.g., monitoring) leading to greater variability in steering response, while active tasks (e.g., gaming) had a more pronounced effect on lane-keeping performance. Notably, moderate NDRT engagement partially offset fatigue-induced lateral control degradation, though fatigue still adversely impacted steering stability without affecting RTOT. However, active NDRTs (e.g., gaming) disrupted vehicle control (in terms of lane keeping) more severely than fatigue alone, underscoring the need for balanced workload management in automated driving systems.

These results challenge assumptions about age-related declines in TOP, emphasizing the role of cognitive functions, driving experience, and adaptive strategies. Future research should develop personalized automation strategies that adapt TOR parameters – such as TB duration and modality – based on individual driver state, ensuring both safety and efficiency in semi-autonomous driving.

Abbreviations

NDRT	Non driving related task
SAE	Society of Automotive Engineers
L3	Level 3 autonomous vehicles as per SAE classification of AVs
AV	Autonomous Vehicles
TOR	Takeover Request
TOP	Takeover Performance
TB	Time Budget
RTOT	Reaction Takeover time
HMI	Human-Machine Interface
RT	Reaction Time
TOT	Takeover Time
TTC	Time to Collision
YD	Younger driver
MD	Middle aged drivers
OD	Older drivers
SWA	Steering Wheel Angle
SWS	Steering Wheel Speed
SDLnP	Standard Deviation in Lane Position

Table of Contents

Older Driver Behaviour in Level 3 Autonomous Vehicles	1
DECLARATION	2
Summary	3
Acknowledgement	4
PUBLICATIONS ARISING FROM THE PRESENT WORK	5
Abstract.....	6
Abbreviations.....	7
Table of Contents	8
List of Tables	13
List of Figures	14
Chapter 1 Literature Review of Older Drivers’ Performance in Level 3 Autonomous Vehicles	17
1.1 Introduction	17
1.1.1 Autonomous Vehicles	18
1.1.2 Takeover Request	19
1.1.3 Factors Affecting Takeover Request.....	19
1.2 What is Mental Workload?	21
1.2.1 Multiple Resource Theory	21
1.2.2 Malleable Attention Theory.....	21
1.3 Methods	22
1.3.1 Problem Statement	22
1.3.2 Search Strategy.....	22
1.3.3 Scope of the review and Inclusion/ Exclusion criteria	23
1.4 Main Findings	25
1.4.1 Takeover Performance Measures.....	27

1.5 Discussion	28
1.5.1 Demographic Impact on NDRTs	28
1.5.2 TOR Configuration (Time Budget) and Demographics	29
1.5.3 Driver’s Cognitive and Psychological State and Demographics	29
1.6 Conclusion.....	29
1.7 Research Questions	30
1.8 Thesis Structure.....	30
Chapter 2 An Investigation of Executive Function in Younger and Older Adults – Implications for Takeover Request Functionality in Autonomous Vehicles	32
2.1 Introduction.....	32
2.1.1 Problem Statement	33
2.1.2 Study Objectives	33
2.2 Conceptual Framework: Linking EF to Takeover Requests.....	33
2.2.1 Executive Functions and Takeover Request.....	33
2.3 Experimental Design: Investigating Age-Related Effects on EF	36
2.3.1 Description of Takeover Tasks	36
2.4 Research Questions	40
2.5 Method	41
2.5.1 Design	41
2.5.2 Participants	41
2.5.3 Materials.....	42
2.5.4 Measures	42
2.5.5 Experimental Procedure	43
2.6 Results	45
2.6.1 Experiment One.....	45
2.6.2 Experiment Two	47
2.7 Discussion	50
2.7.1 Theoretical Implications – Conceptual Framework	50

2.7.2 Experimental Implications.....	51
2.8 Conclusion.....	58
2.9 Implications for L3 AVs	59
2.10 Limitations of the Current Study.....	59
Chapter 3 Age Differences in Takeover Performance Across Varying Time Budgets in Level 3 Autonomous Vehicles	61
3.1 Introduction	61
3.1.1 Effect of Age on Takeover Performance (TOP)	61
3.1.2 Effect of TB on TOP.....	62
3.2 Method	63
3.2.1 Experiment Design.....	63
3.2.2 Experiment One.....	64
3.2.3 Experiment two	67
3.3 Results	69
3.3.1 Experiment One.....	70
3.3.2 Experiment two	72
3.4 General Discussion.....	75
3.4.1 Effect of Age on Takeover Performance	75
3.4.2 Effect of Time Budget on Takeover Performance	76
3.5 Conclusion.....	78
3.6 Limitations	79
Chapter 4 Effect of NDRT on Takeover Performance	80
4.1 Introduction	80
4.1.1 Effect of NDRT	80
4.1.2 Takeover Performance.....	81
4.2 Method	82
4.2.1 Design.....	82
4.2.2 Participants	84

4.2.3 Materials.....	85
4.2.4 Procedure.....	85
4.3 Results	87
4.3.1 Behavioural Measures	88
4.3.2 Qualitative Measures.....	91
4.4 Discussion	92
4.4.1 Active Engagement in NDRTs (Game & Video) will have a significant effect on takeover performance than “passive” engagement in the NDRT (Observation).....	92
4.5 Conclusion.....	95
Chapter 5 Age Differences in Fatigue Development and Perception in Level 3 Autonomous Vehicles.....	97
5.1 Introduction	97
5.1.1 Fatigue in Automated Vehicles.....	98
5.1.2 Research Questions	100
5.2 Methodology	100
5.2.1 Participants	100
5.2.2 Materials.....	101
5.2.3 Experiment Design.....	102
5.2.4 Experimental Condition	102
5.2.5 Takeover Performance.....	103
5.2.6 Procedure.....	104
5.3 Results	106
5.3.1 Descriptive Statistics	106
5.3.2 Statistical Analysis	106
5.4 Discussion	111
5.4.1 Effect of Age on Takeover Performance	111
5.4.2 Effect of Task Conditions (Drive monitoring Vs NDRT) on Takeover Performance.....	112
5.4.3 Fatigue Induction and Mitigation Using Workload from NDRTs	114
5.5 Conclusion.....	115

5.6 Limitations, Future Directions and Practical Implications	116
Chapter 6 Discussion.....	117
6.1 Summary of Previous Chapters	117
6.2 Methodological Considerations.....	118
6.3 Empirical Contributions	120
6.3.1 Effect of Age on TOP	120
6.3.2 Effect of Workload	121
6.4 Theoretical contributions.....	125
6.4.1 Multiple Resource Theory (MRT), Cognitive Control Hypothesis (CCH) and Takeover Performance.....	125
6.4.2 Compensatory Control Model and Takeover Performance	126
6.4.3 Proposed Framework.....	126
6.5 Conclusion and Limitations.....	128
6.6 Future Research.....	129
6.6.1 Proposed Methodology.....	129
6.6.2 Expected Contributions and Impact	130
6.7 Key Findings	130
6.7.1 Empirical Findings	130
6.7.2 Theoretical Contributions	131
References	132
Appendix	157
Appendix A	157
Appendix: A1	159

List of Tables

Table 1. 1. Shows the Taxonomy and definition of levels of automation as per SAE.....	19
Table 1. 2. Review of selected studies.....	25
Table 2. 1. Participants' demographics and educational background for Experiment 1.	45
Table 3. 1. Participants' Demographics for Experiment 1...	64
Table 3. 2. Participants' demographics for Experiment two.....	67
Table 3. 3. Describe a Latin Square experiment design where H stands for High TB (5 seconds) and L stands for Low TB (2.5 seconds).	69
Table 3. 4. Descriptive Statistics and p-value for TOP metrics based on TB as a factor for Experiment 1.....	70
Table 4. 1. TOP metrics and their definition.	84
Table 4. 2. Participants' demographics.	84
Table 4. 3. ANOVA Table for the effect of NDRT and Task Iterations on TOP metrics (only the significant metrics are reported).....	89
Table 5. 1. Participants' demographics.	100
Table 5. 2. Takeover Performance metrics.	104
Table 5. 3. Descriptive statistics of each of the TOP metrics.	106

List of Figures

Figure 1. 1. PRISMA search strategy.	24
Figure 1. 2. Time Budget as adopted in different studies.	26
Figure 1. 3. Response Takeover Time (RTOT) from different studies.	27
Figure 2. 1. Depicting different elements of executive functioning and attention involved during a TOR scenario.	35
Figure 2. 2. A snapshot depicting (a) Updating Task, (b) Shifting Task, (c) Inhibition Task during the experimental run of these tasks, (d) Combined Task, and (e) describing the appearance of “Target” image along with “Distractor” images.	38
Figure 2. 3. Showing the sequence of the tasks presented in experiment one (above one). (P2) the sequence of tasks presented to the participants in experiment two.	43
Figure 2. 4. Mean reaction time (in seconds) for ‘Over 60’ years old and ‘18-35’ years old adults across each tasks i.e. Updating, Shifting, Inhibition, and combined tasks. The error bars here represent 95 % confidence interval.	46
Figure 2. 5. Plot of the number of incorrect responses for each task across both age groups.	47
Figure 2. 6. Mean Reaction time for each task as observed in Experiment two.	48
Figure 2. 7. Plot of Z-scored normalized data for Reaction Time, Cumulative Errors Vs Task Difficulty.	49
Figure 2. 8. Frequency plot of incorrect responses by participants for each task in the Experiment two and by only younger participants from the Experiment One. The tasks are named in the order of their presentation.	49
Figure 2. 9. Mean Perceived Task difficulty scores across task conditions in the presentation order.	50
Figure 3. 1. Driving Simulator along with the mounted Tablet used in this study.	65
Figure 3. 2. (a) The experimental driving scenario. (b) the schematic diagram of the TOR events in the experiment.	66
Figure 3. 3. The Driving Simulator used in Experiment two.	68
Figure 3. 4. (a) Steering Wheel angle for different TB and (b) Steering Wheel Speed for different TB.	71
Figure 3. 5.(a) Reaction Takeover Time for different TB and (b) Lateral SDLnP for different TB.	71

Figure 3. 6. (a) Vehicle speed for older and younger drivers and (b) Reaction Takeover Time for older and younger drivers.....	72
Figure 3. 7. Driving behaviour statistics after TOR (a) SWA, (b) SWS, (c) RTOT, and (d) SDLnP. Low refers to 2.5 seconds of TB and High refers to 5 seconds of TB. ** significant at $p < 0.00574$	
Figure 3. 8. NASA TLX for High Vs Low TB.....	75
Figure 4. 1. Driving Simulator at Trinity College Dublin with a mounted Samsung Tablet.....	85
Figure 4. 2. (a) The experimental driving scenario. (b) The time points for each phase of the driving scenario. Adapted from Zhang et al. (2022). (c) NDRT tasks as Task iterations which were randomized during the experiment.....	86
Figure 4. 3. Gendered driving metrics after TOR: (a) SWA, (b) RTOT, (c) SDLnP, and (d) Deceleration	88
Figure 4. 4. SWA metric plot for task condition and iteration.....	89
Figure 4. 5. SDLnP metric plot for task condition and iteration.	90
Figure 4. 6. RTOT metric plot for task condition and iteration.....	90
Figure 4. 7. Deceleration metric plot for task condition and iteration.	91
Figure 4. 8. Gender differences in: (a) subjective perception of NASA-TLX workload and (b) attitudes towards AVs.....	92
Figure 5. 1. Demographics distribution of Participants.	101
Figure 5. 2. High Fidelity dynamic driving simulator at Tobii.....	101
Figure 5. 3. Driving virtual environment of the simulator from SCANeR software.....	102
Figure 5. 4. Longitudinal and Lateral Deceleration for Fatigue and NDRT conditions. ** significance at $p < 0.05$	107
Figure 5. 5. Longitudinal and Lateral SDLnP for Fatigue and NDRT conditions. ** significance at $p < 0.05$	108
Figure 5. 6. Steering Wheel Angle and Steering Wheel Angle Speed for Drive Monitoring and NDRT condition. ** significance at $p < 0.05$	108
Figure 5. 7. Reaction Takeover Time and Brake for Fatigue and NDRT conditions.....	109
Figure 5. 8. Frequency distribution of reported KSS levels with their corresponding description.....	110
Figure 5. 9. Self-reports for Younger and Older adults in different conditions for (a) KSS and (b) NASA-TLX. ** significant for $p < 0.05$	110

Figure 5. 10. Self-reported (a) KSS for NDRT and Fatigue conditions, and (b) NASA-TLX for fatigue and NDRT condition. ** significant for $p < 0.05$	111
Figure 6. 1. Representing the factors affecting TOP and its quantified metrics.	119
Figure 6. 2. Conceptual Framework detailing the effect of changes in workload on performance and driving output.	127

Chapter 1 Literature Review of Older Drivers' Performance in Level 3 Autonomous Vehicles

1.1 Introduction

As the global population ages, the number of older drivers on the road is steadily increasing. According to the World Health Organization (WHO), the proportion of people aged 60 and over is expected to double by 2050, rising from 12% to 22% of the global population (WHO, 2021). This demographic shift has significant implications for road safety, as older drivers face unique challenges that influence their driving performance.

Older drivers are overrepresented in specific types of collisions, particularly those occurring at intersections (Preusser et al., 1998). Studies have shown that drivers aged 65 and older are more likely to be involved in crashes involving right turns, failure to yield, and misjudgement of gaps in traffic (IIHS, 2020). These incidents are often attributed to age-related declines in cognitive and physical abilities. For instance, reduced visual acuity, slower reaction times, and diminished peripheral vision can impair an older driver's ability to accurately assess oncoming traffic or judge distances (NHTSA., 2019). Additionally, cognitive impairments, such as difficulty processing complex information or multitasking, can lead to errors in decision-making at intersections (Anstey et al., 2005). These factors contribute to their overinvolvement in intersection-related crashes, which are among the most common and dangerous types of collisions for older drivers.

Another area of overinvolvement is single-vehicle crashes, such as running off the road or hitting stationary objects. These incidents may result from age-related declines in motor skills, such as reduced muscle strength and flexibility, which can affect steering and braking control (NHTSA, 2019). Furthermore, older drivers are more susceptible to medical conditions like arthritis, Parkinson's disease, or stroke, which can impair their ability to operate a vehicle safely (Carr et al., 2006). Medications used to manage these conditions may also have side effects, such as drowsiness or dizziness, further increasing the risk of crashes (IIHS, 2020). Additionally, older drivers experience age-related frailty, making them more likely to sustain serious injuries from collisions that might cause minimal harm to younger individuals. This phenomenon, known as "frailty bias," complicates the analysis of collision data, as most collisions are only included in official databases if they result in serious injury or involve an insurance claim. Consequently, older drivers' crash rates may appear disproportionately high, not solely due to their driving behaviour but also because of their increased vulnerability to injury.

To conduct a fair risk factor analysis across age groups, it is essential to account for these biases. Comparisons should be based on metrics such as yearly mileage exposure or other measures that

encapsulate older drivers' compensatory behaviours, such as avoiding high-risk driving conditions (e.g., nighttime driving or heavy traffic) (Molnar, 2003). Older drivers often self-regulate their driving habits to mitigate risks, which can lead to under-involvement in certain types of collisions, such as high-speed crashes or those involving alcohol impairment. They are less likely to engage in risky behaviours like speeding, aggressive driving, or driving under the influence of alcohol (NHTSA, 2019). While this cautious approach reduces their likelihood of being involved in certain types of collisions, it may also limit their mobility and independence, potentially leading to social isolation and reduced quality of life (Chihuri et al., 2016).

Despite these patterns, it is important to recognize that not all older drivers face the same challenges. Individual differences in health, driving experience, and adaptability play a significant role in determining driving performance (Anstey et al., 2012). Some older drivers remain highly competent and safe, while others may struggle with age-related declines (Li S. et al., 2021). Addressing these challenges requires a multifaceted approach, including regular assessments of older drivers' physical and cognitive abilities, improved road design to accommodate their needs, and educational programmes to promote safe driving practices (Carr et al., 2010). By understanding the unique challenges faced by older drivers and accounting for biases such as frailty bias, society can work towards creating a safer and more inclusive driving environment for all.

Level 3 autonomous vehicles, in theory, have the potential to solve these mobility issues of older drivers by acting as a co-pilot that will work with the driver and support them in difficult driving situations and manoeuvres. These vehicles could aid the older drivers in high congestion scenarios such as intersections where they are already over-represented in crashes. This study in the following sections will discuss whether autonomous vehicles could be assistive to older drivers.

1.1.1 Autonomous Vehicles

Autonomous vehicles (AVs) are evolving at a very rapid pace, however, it will still take a couple of years before we can expect a complete self-driving vehicle (Hancock et al., 2019). AVs are defined as vehicles that will be able to move safely on roads by sensing the road environment with little to no human involvement. According to the Society of Automotive Engineers (SAE), AVs are categorized into six levels, ranging from level 0 to level 5; the details about each level have been explained in Table 1.1 (SAE, 2021). It is expected that in the near future, level 3 and level 4 vehicles are going to dominate the AV market (Li et al., 2019). These vehicles provide the driver with the option to engage in non-driving related tasks (NDRTs) such as texting, watching a video, eating, talking with a passenger, etc. These vehicles could, at least in theory, help older drivers to continue to drive longer and safer (Fatima et al., 2020; Gormley & O'Neill, 2019). Most of the research that involves older drivers and autonomous vehicles is in the form of surveys or review articles (Faber &

van Lierop, 2020). However, thorough empirical research is required to substantiate these claims and to understand their outreach, benefits, and shortcomings as well.

1.1.2 Takeover Request

Level 3 or conditional automation – even with their limited automation – enables the driver to engage in non-driving related tasks (NDRT). They have limited automation in the sense that under very specific situations such as construction zones, school zones, broken vehicles ahead, etc., the driver may have to take back control of the vehicle. The system will notify the driver of these situations by initiating a ‘Takeover request (TOR)’ to the driver to take control of the vehicle – which could be a visual, aural, tactile or a combination of these. The system will also provide a Time Budget (TB) to the driver within which they have to take back control of the vehicle. After the notification of the takeover request, the driver has to shift their attention to driving and gain sufficient situational awareness so that they make a smooth and safe transition (Gold et al., 2013). The time it takes the driver to assume control in response to a TOR is called Response takeover time (RTOT). This TOR task has been a focus of much interest in the research community with studies looking into the modality of the takeover request (verbal, visual, or tactile), takeover time, the cognitive load, the psychological state of the driver during automation, situational awareness, engagement of the driver

Table 1. 1. Shows the Taxonomy and definition of levels of automation as per SAE.

Level	Name	SAE definition
0	No automation	The lowest level of automation in SAE levels of automation. It does not perform any form of lateral or longitudinal control, with limited assistance in terms of night vision, navigation system, or lane departure warnings.
1	Driver assistance	It provides lateral and longitudinal control of the vehicle but not both simultaneously.
2	Partial automation	This driving automation can perform both lateral as well as longitudinal control at the same time. However, the driver is responsible for the safety of the vehicle and thus monitor the vehicle.
3	Conditional driving automation	It provides all the functionalities of level 2 automation with the addition that the driver may completely disengage from but needs to be available to assume controls.
4	High driving automation	This level of automation can drive completely without human intervention within certain areas, however, it may require human presence to operate the vehicle outside those areas.
5	Full driving automation	This level is fully automated in the truest sense as it does not require human presence at all and can drive under all conditions.

to music or chatting with a passenger (Gold et al., 2016). This highlights the need for autonomous systems to account for the varying levels of driver distraction when issuing TORs.

TOR configurations, particularly the time budget (TB) and modality of the TOR, also play a pivotal role in shaping driver responses. A shorter time budget – such as 5 seconds or less – places greater cognitive demands on drivers, increasing the likelihood of errors during the transition (Eriksson et al., 2017). On the other hand, longer time budgets allow drivers more opportunity to reorient themselves and regain control. Additionally, the modality of the TOR – whether it is visual, auditory, or haptic – can significantly influence how quickly and effectively drivers respond. Research suggests that multimodal alerts, which combine visual and auditory cues, are more effective at capturing attention and improving response times compared to single-modality alerts (Petermeijer et al., 2017). This underscores the importance of designing intuitive and attention-grabbing TOR systems.

The driver's cognitive and psychological state is another critical factor. Fatigue, stress, or distraction can impair a driver's ability to process information and execute timely responses to a TOR (Körber et al., 2016). For instance, a fatigued driver may take longer to disengage from an NDRT and reorient to the driving task, increasing the risk of errors. Additionally, cognitive overload or excessive mental workload (MWL) can hinder decision-making and slow reaction times, particularly in high-demand driving scenarios. Conversely, when cognitive demand is balanced – neither too high nor too low – drivers are more likely to perform well during transitions. This highlights the need for systems that can monitor driver states, detect fluctuations in MWL, and adapt TOR strategies accordingly.

Demographic factors, such as age and gender, also influence TOR performance. Older drivers, for example, may experience slower reaction times and reduced cognitive flexibility, making it more challenging to respond to TORs (Gold et al., 2016). Additionally, older drivers may experience a greater cognitive workload when switching from passive monitoring to active control, particularly if they are required to process complex situational information in a short timeframe (Peng et al., 2022). Age-related declines in vision, hearing, or motor skills can further complicate the transition process. Gender differences have also been observed, with some studies suggesting that males may respond faster to TORs than females, though the underlying reasons for this disparity remain unclear and warrant further investigation (Eriksson & Stanton, 2017).

In conclusion, factors such as NDRTs, vehicle configurations, drivers' cognitive state & mental workload, and demographics significantly influence the effectiveness of TORs. Addressing these factors through improved system design, adaptive interfaces, and targeted driver training is essential for ensuring safe and efficient transitions in autonomous driving. By tailoring TOR systems to account for these human factors, we can enhance the safety and reliability of autonomous vehicles, paving the way for their widespread adoption.

1.2 What is Mental Workload?

Arguably one of the biggest impacts of cognitive functioning while driving is mental workload, which is a multifaceted construct that has been widely studied in fields such as Human Factors, Ergonomics, Systems Design, and Neuro-ergonomics. It refers to the total cognitive demand placed on an individual to accomplish a specific task within a given timeframe (Young et al., 2015). Mental workload encompasses various cognitive demands, including attention, memory, and decision-making, as well as the subjective effort experienced by an operator to meet those demands. The level of workload an individual experiences can significantly impact task performance, stress levels, and overall job satisfaction, particularly in high-stakes environments such as driving, aviation, and medicine (Cain, 2007; de Winter et al., 2014).

In the context of Level 3 automated vehicles, factors such as NDRTs, TB, and fatigue contribute to the mental workload experienced by the driver during a TOR. These factors influence takeover performance (TOP) by modulating cognitive effort and resource allocation, making the study of workload essential for improving safety and efficiency in automated driving systems (Eriksson & Stanton, 2017). Several theoretical models have been proposed to explain how mental workload operates and affects human performance. Below are some key theories relevant to the study of workload in the driving domain.

1.2.1 Multiple Resource Theory

Multiple Resource Theory (MRT), proposed by Wickens (1984) and further expanded by de Waard (1996), posits that human cognitive resources are divided into different categories based on input modality (e.g., visual, auditory), processing codes (e.g., spatial, verbal), and response execution mechanisms. According to this theory, if multiple tasks require resources from the same cognitive channel, the mental workload increases, leading to reduced performance and increased cognitive strain. However, if tasks rely on separate resource pools, they can be performed concurrently with minimal interference. In the context of automated driving, MRT helps explain why simultaneous engagement in visual-manual NDRTs, such as texting, significantly impairs takeover performance, whereas auditory-based tasks may have less impact on driving performance (Wickens, 2008).

1.2.2 Malleable Attention Theory

The Malleable Attention Theory (MAT) suggests that attentional resources are not static but can be dynamically adjusted depending on task demands (Young & Stanton, 2002). Under this framework, attentional allocation is influenced by both intrinsic factors (e.g., cognitive capacity, fatigue) and extrinsic factors (e.g., task complexity, urgency). In driving automation, this theory explains how drivers' attentional resources shift between NDRTs and the monitoring of the driving environment. If an individual has been deeply engaged in an NDRT, their attentional resources may be slow to

reallocate to the driving task, leading to delayed reactions during a takeover event (Louw et al., 2015). Conversely, shorter or intermittent engagement in NDRTs may allow for a more efficient redistribution of attention when a TOR is issued.

Understanding these theoretical models is critical for designing effective human-machine interfaces (HMIs) and automation strategies that mitigate excessive workload while ensuring optimal driver performance during takeovers. Future research should focus on refining workload assessment methods and incorporating physiological and subjective measures to develop adaptive automation systems that can dynamically modulate workload demands.

1.3 Methods

This review is based on the PRISMA method – Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Moher et al., 2009). Per the stated method, the problem statement was defined, followed by the scope of this study, and inclusion/exclusion criteria were described. After outlining these fundamental terms of this systematic review, the relevant literature was screened and identified by scanning their titles and abstract, the screened literature then was either included or excluded depending on the inclusion/exclusion criteria. Finally, the included papers were studied thoroughly and taken forward for further review to ascertain if they are relevant to the problem statement defined below.

1.3.1 Problem Statement

“What is the influence of age or age-related declines on the TOP in a level 3 autonomous vehicle and what are the factors that interact with age to affect TOP?”

This research question was developed based on the statement of the PRISMA and the guidelines of PICOS – Population, Intervention, Comparator, Outcomes and Study design (Moher et al., 2009). The purpose of this study is to collate and report studies conducted on older drivers in automated vehicles and identify possible future directions based on their reported results.

1.3.2 Search Strategy

A comprehensive phrase search was conducted using the following keywords: 'older drivers' OR 'older adults' or 'elderly' AND 'takeover request' OR 'takeover' OR 'take over' OR 'transition of control' AND 'level 3 automation' OR 'conditional automation' OR 'conditional autonomous vehicles' OR 'autonomous vehicles' OR 'automated vehicles'. The search was done in April 2021.

These search keywords were used to look for relevant studies in these databases: PubMed, Scopus, Web of Science, ACM digital library, APA PsycInfo, IEEE Xplore, Google Scholar, Science Direct, and EMBASE.

A total of 274 articles were selected initially using the above-mentioned search phrase. After title and abstract screening, further removal of duplicate articles from various search engines non-peer-reviewed conference articles, and book chapters; further exclusion of the papers was done if their full texts were not available, see Figure 1.1. Finally, 38 papers were considered for full review and after carefully reviewing the articles based on their relevance to the abovementioned problem statement, nine papers were included and their findings were reported in this review.

1.3.3 Scope of the review and Inclusion/ Exclusion criteria

1.3.3.1 *Eligibility Criteria*

The studies that were included in this review must fulfil the criteria below:

- The study must be about the transition from automation to manual control of the vehicle; the transition could either be a result of system failure or a planned takeover.
- The study must have words like takeover request, takeover time or transition of control in the text.
- The study must have used age as an independent variable or have studied age as a factor influencing driving performance.
- The study must be about conditional automation (SAE level 3 automated vehicles).
- The study has to be written in English.

The studies which did not have full text available or were review articles or published as book chapters or non-experimental papers such as survey articles, review papers, interview papers or were published in a non-peer review journal or conference were excluded. There were no criteria set on publication date – relevant papers published from any date were accepted.

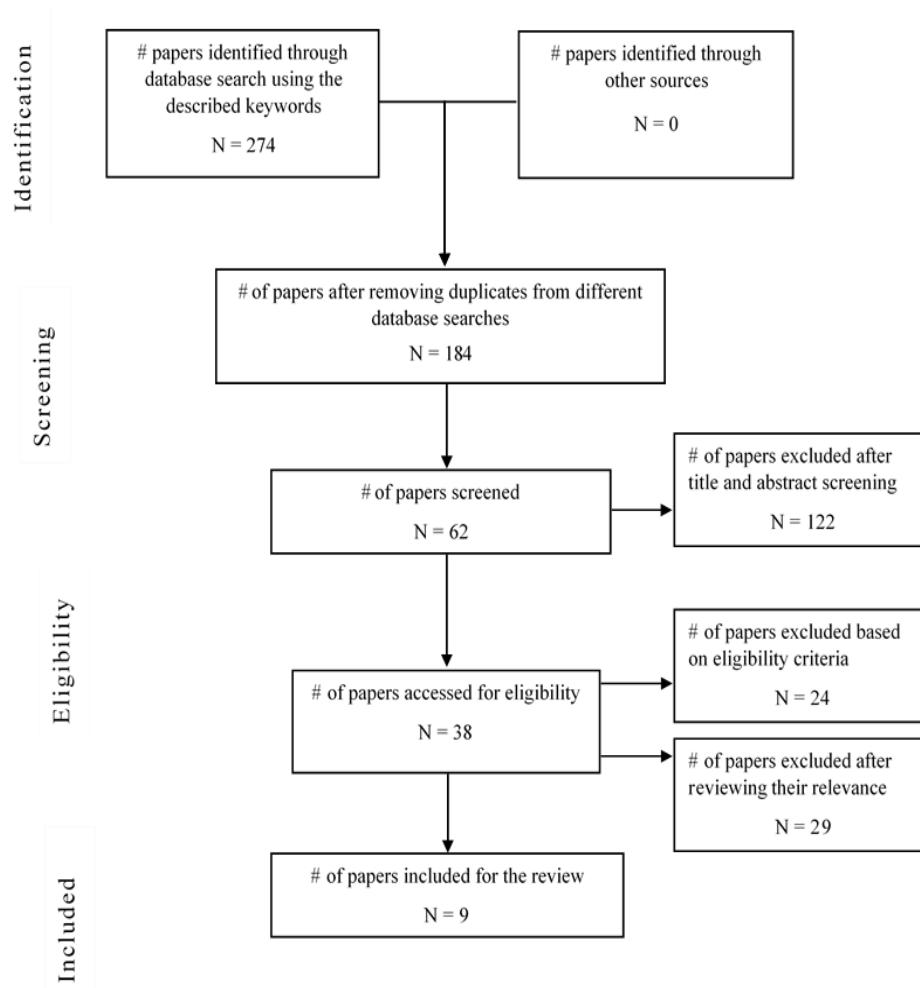


Figure 1. 1. PRISMA search strategy.

1.4 Main Findings

In this section, different findings related to the older drivers' workload and performance issues have been reported in Table 1.2. This table also includes a summary and findings of the reviewed literature from a workload perspective. The key terms used in the Table 1.2 have been defined below.

Time Budget (TB): It is defined as the time provided by the vehicle system at the initiation of a TOR

Table 1. 2. Review of selected studies.

Study Name	Study Participants	NDRT	Workload Observed	Workload Estimated	Time Budget	Response Time (RTOT)
Körber et al., (2016)	YD: 36 (19-28 yrs), OD: 36 (60-79 yrs)	20 Questions (Cognitive Distraction)	Yes (NDRT & Traffic Density)	No	7s	YD: 3.3s (No task), 3.7s (task); OD: 3.4s (No task), 3.2s (task)
Miller et al. (2016)	YD: 12 (15-19 yrs), MD: 12 (17-69 yrs), OD: 12 (70-81 yrs)	Watching a movie, reading a story on a tablet, supervising car's driving	Yes (Engagement in NDRTs)	No	5s	X
Clark & Feng, (2017)	YD: 17, OD: 18	Voluntary choice of participant	Yes (Engagement in NDRTs)	No	7.5s and 4.5s	YD: 2.18s, OD: 1.71s
Li et al. (2018)	YD: 37 (20-35 yrs), OD: 39 (60-81 yrs)	Reading aloud from a mounted device	Yes (Engagement in NDRTs)	No	20s	OD: 4.32s, YD: 3.60s (mean of all weather conditions)
Wu et al. (2019)	115 participants (18-75 yrs), divided into 5 groups of 23 each	No NDRT	Yes (Scheduled manual transition)	No	6s	X
Li et al. (2019a)	YD: 37 (20-35 yrs), OD: 39 (60-81 yrs)	Reading aloud from a mounted device	Yes (Engagement in NDRTs)	Yes (NASA RTLX score)	20s	OD: 3.12s, YD: 3.4s
Li et al. (2019b)	YD: 37 (20-35 yrs), OD: 39 (60-81 yrs)	Reading aloud from a mounted device	Yes (Engagement in NDRTs)	No	20s	OD: 3.38s, YD: 2.84s
Wu et al. (2020)	YD: 12 (19-32 yrs), MD: 12 (37-50 yrs), OD: 12 (56-74 yrs)	Watching a video (30 min) OR Watching a video (10 min) + Playing game (10 min) + Watching video (10 min)	Yes (Engagement in NDRTs)	No	6s	X
Huang and Pitts (2020)	YD: 24 (20-24 yrs), OD: 24 (66-76 yrs)	Watching a video	No	No	X	X

YD: Younger Drivers; OD: Older drivers; MD: Middle age drivers

within which the driver has to assume control of the vehicle (Gold et al., 2013). The TB should be

sufficient for the driver to disengage from the NDRT, gain sufficient situational awareness, and assume vehicle control. Figure 1.2 shows the time budgets (TB) selected in the reviewed studies, with most opting for less than 10 seconds. TBs under 5 seconds are classified as urgent or critical takeovers (Doubek et al., 2020). Notably, Shuo Li (2019), used a longer TB of 20 seconds to account for age-related cognitive decline and reduce stress in older drivers during critical takeovers.

The issue of the influence of TB on takeover performance has yet to be understood clearly. Studies have reported contradictory findings on this issue. A few studies reported TB below 5 seconds is not sufficient for a safe takeover (Beukel & Voort, 2013), while others found this TB to be sufficient (Clark & Feng, 2017; B. Mok et al., 2015). A study by Gold et al. (2013) found that shorter TB results in quicker takeover, however, the takeover quality is affected negatively.

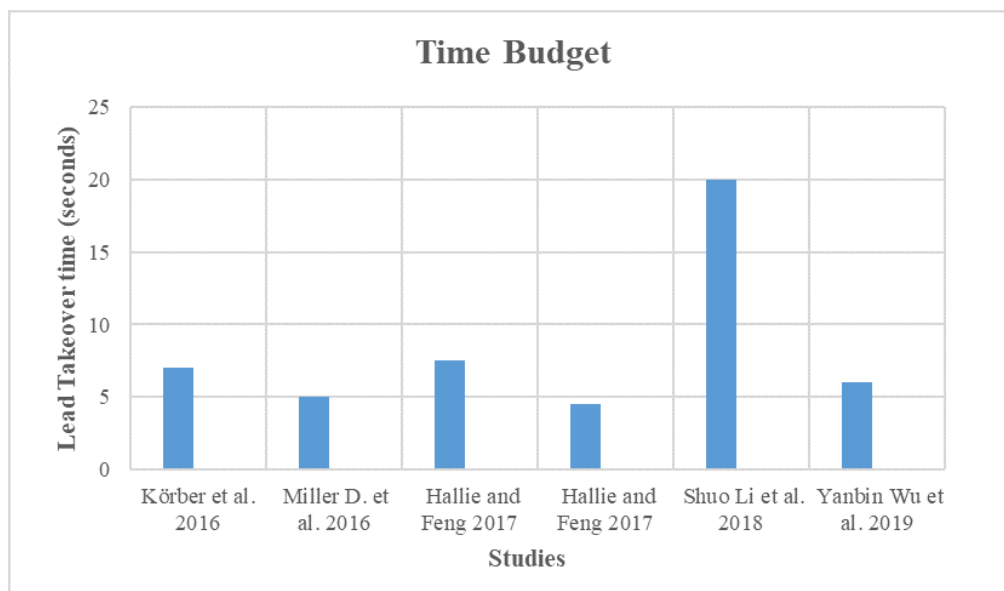


Figure 1. 2. Time Budget as adopted in different studies.

Response takeover time (RTOT): It is defined as the time it takes the driver to make the conscious input in response to the takeover request. The conscious input has been defined as the manoeuvre of the steering wheel angle of 2° and/or 10% movement of accelerator or brake pedal positions (B. K. Mok et al., 2015). Figure 1.3 shows a comparison of response takeover times between older and younger drivers as reported in the studies selected for this review. An interesting observation from figure 1.3 is that for both older and younger drivers the response takeover time lies below 5 seconds, which suggests that drivers can regain vehicle control fairly quickly after a TOR, irrespective of age. Also, it can be observed that in most of the studies conducted on older drivers' performance in L3 automation, the performance (in terms of takeover time) is almost comparable to younger drivers, if not better in some studies. The only exception reported is by Li et al. (2018), where there is an observable difference in performance between the two groups. The reason reported by the authors for this apparent difference in performance is that they have adopted a longer TB of 20 seconds. The

older drivers, being cautious and slow, used this TB to assume vehicle control at their own pace. This led to an increase in their response takeover time, but their takeover quality improved significantly.

1.4.1 Takeover Performance Measures

A range of takeover performance measures have been adopted by different studies to quantify takeover performance and takeover quality. Takeover performance is defined as the performance of a driver during a takeover scenario, which means how quickly the driver responded to the takeover request. Usually, takeover performance is quantified using time-based measures such as Takeover time or Reaction takeover time, Indicator time, lane deviation, brake inputs, etc. While the takeover

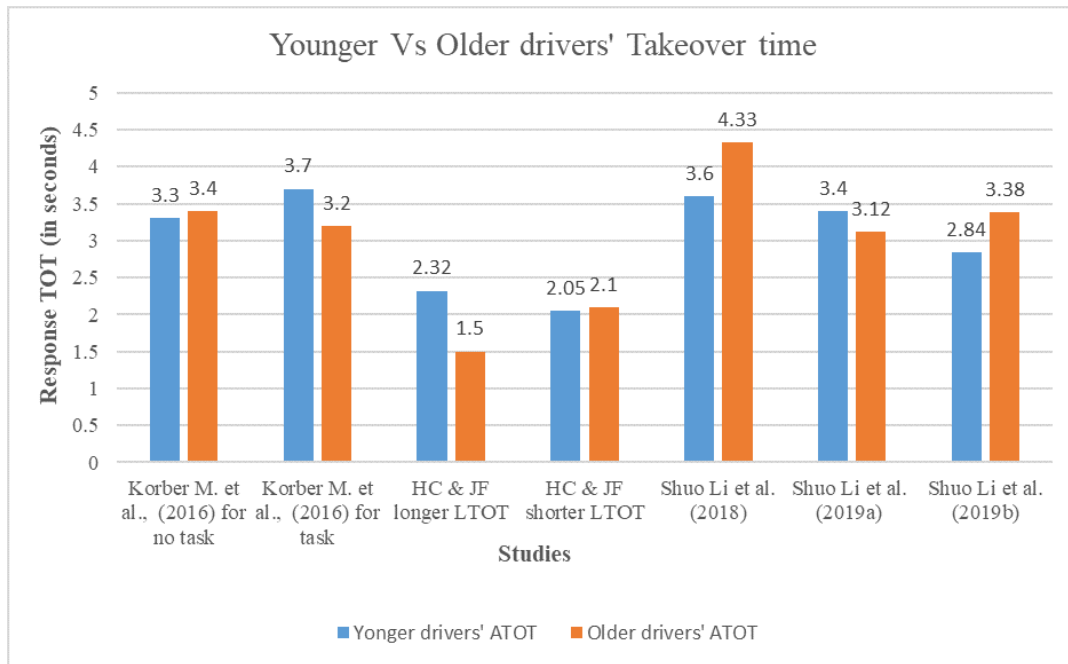


Figure 1. 3. Response Takeover Time (RTOT) from different studies.

quality is defined in terms of the safety and conservative aspect of the takeover. The measures commonly used to quantify takeover performance are steering wheel angle, steering wheel reversal rate, minimum time to collision, lateral and longitudinal acceleration/deceleration, collisions and critical encounters, etc. A select few most commonly used measures have been defined further for better clarity and understanding.

Response Takeover Time (RTOT): It is the time it takes the driver after the initiation of TOR to their conscious input to the vehicle, which is defined in literature as a manoeuvre of the steering wheel angle of 2° and/or 10% movement of accelerator or brake pedal positions. It is also used as independently as Time-steer-1 and Time-Brake-0.1 in a study by Wu et al. (2019).

sdSteer: It refers to the standard deviation of the steering wheel angle. It is a measure of the smoothness of the manoeuvre.

Minimum Time-to-Collision (TTC): It is an effective measure in assessing the severity of potential collisions. It refers to the time required for the test vehicle to collide with the stationary vehicle obstructing the driving lane if it continues at its speed at the time it changes to the next lane completely.

Steering Wheel Angle (SWA): It represents a change in angle in degree from the centre position of the steering wheel.

1.5 Discussion

The discussion on the impact of NDRTs, TOR configurations, and drivers' cognitive and psychological states on takeover performance in automated vehicles reveals a nuanced interplay of factors, particularly when considering demographic characteristics such as age and gender. This narrative synthesizes findings from multiple studies to provide a comprehensive understanding of these dynamics, enriched with additional research to contextualize and expand upon the existing discussion.

1.5.1 Demographic Impact on NDRTs

The impact of NDRTs on takeover performance is closely tied to age. Körber et al. (2016) found that while there were no significant differences in takeover performance across age groups during NDRT engagement, older drivers exhibited more cautious behaviours, such as harder braking, as noted by Clark and Feng (2017). This cautious behaviour may reflect a compensatory mechanism to offset age-related declines in cognitive and motor functions. Li et al. (2018) further highlighted that older drivers took longer to react and exhibited compensatory behaviours like harder braking and acceleration during NDRT engagement, which aligns with findings by Reimer et al. (2013) that older adults often adopt conservative strategies to manage cognitive load.

Interestingly, Wu et al. (2020) found that while younger drivers benefited from NDRTs as a countermeasure to drowsiness, older drivers did not experience similar benefits. Instead, their performance declined, suggesting that NDRTs may impose additional cognitive demands on older drivers. This is consistent with Verhaeghen et al. (2003) who argued that older adults are more susceptible to cognitive overload when multitasking, which can exacerbate performance deficits during critical driving tasks.

Gender differences may further compound this issue. Gwyther and Holland (2012) found that older female drivers, in particular, tend to exhibit more cautious driving behaviours, which may extend to their engagement with NDRTs. For instance, older women may be more likely to disengage from secondary tasks earlier than older men when a takeover request is imminent, prioritizing safety over task completion.

1.5.2 TOR Configuration (Time Budget) and Demographics

The configuration of TORs, particularly the TB, is also influenced by age. Clark and Feng (2017) found that older drivers performed comparably to younger drivers across short (4.5s) and long (7.5s) TBs but preferred longer TBs, especially when engaged in secondary tasks. This preference aligns with research by Salthouse (1996), which shows that older adults require more time to process information and execute responses due to age-related declines in processing speed. Li et al. (2019a) extended this research by examining a significantly longer TB of 20 seconds, revealing that older drivers took more time to regain control but were able to compensate for age-related declines, ultimately improving takeover quality. Huang and Pitts (2020) also found that both younger and older drivers preferred multimodal alerts (combining aural, visual, and tactile signals), but older drivers allocated more attention to the road, indicating differences in attentional strategies.

This aligns with findings by Eriksson et al. (2017), who emphasized that multimodal alerts can enhance situational awareness, particularly for older drivers who may rely more on visual cues due to age-related declines in auditory processing.

1.5.3 Driver's Cognitive and Psychological State and Demographics

The cognitive and psychological state of drivers, particularly fatigue and drowsiness, is deeply intertwined with age and gender. Wu et al. (2019) investigated scheduled automation to manual transitions as a countermeasure to drowsiness and found that while these transitions temporarily alleviated drowsiness, their effects were short-lived, lasting only about 10 minutes. Older drivers were particularly negatively impacted, exhibiting slower responses in steering and braking during critical events.

This suggests that scheduled manual transitions may impose unnecessary cognitive demands on older drivers, a finding supported by Hess and Knight (2021) who argued that older adults are less able to recover from cognitive fatigue.

Wu et al. (2020) found that while younger drivers benefited from NDRTs as a countermeasure to drowsiness, older drivers did not, and their performance suffered as a result. This discrepancy may be attributed to age-related differences in cognitive resilience, as highlighted by Staub et al. (2013), who found that older adults are less able to maintain performance under conditions of sustained cognitive load. Older women, in particular, may require tailored interventions to manage fatigue, such as more frequent breaks or personalized drowsiness detection systems.

1.6 Conclusion

Overall, the research underscores the importance of tailoring automated vehicle systems to the needs of older drivers. Longer TBs, multimodal alerts, and age-friendly HMIs can mitigate age-related declines in takeover performance. However, the cognitive and psychological states of drivers,

particularly older adults, must be carefully considered when designing interventions to manage fatigue and drowsiness. Future research should continue to explore the interplay between demographic characteristics, cognitive load, and system design to enhance the safety and usability of automated vehicles for all drivers.

1.7 Research Questions

By examining these factors, this research thesis aims to uncover how older drivers respond to takeover requests and how automated driving systems can be optimized to accommodate their unique needs. This holistic approach not only advances our understanding of the human factors involved in automated driving but also provides actionable insights for designing inclusive and safe systems that enhance driver performance and confidence across diverse user groups.

1. *What are the fundamental cognitive functions required to perform a takeover request scenario effectively, and how do these functions vary across different age groups?*
2. *What is the effect of Time Budget on Takeover Performance, and how does age influence this relationship?*
3. *How does engagement in an NDRT during automated driving affect takeover performance?*
4. *What role does fatigue or drowsiness play in takeover performance, and how does age influence the effectiveness of countermeasures such as NDRTs or scheduled manual transitions?*

By addressing these questions, this thesis can contribute to the development of automated driving systems that are not only technologically advanced but are also inclusive and safe for more diverse user groups.

1.8 Thesis Structure

The following sections have tried to unwrap key concepts such as driving, older adults, autonomous vehicles, cognitive functions and ageing, and how these are linked to driving and as such to takeover request functionality.

Chapter 1 reviews essential concepts and situates them within ongoing research, providing a systematic literature review on Level 3 takeover requests and older drivers.

Chapter 2 presents findings on executive functions, ageing, and their implications for takeover request performance.

Chapter 3 examines the interaction between TB and ageing and their impact on takeover performance.

Chapter 4 investigates the influence of NDRTs on takeover performance.

Chapter 5 analyses the effects of higher levels of automation on fatigue and its impact on takeover performance, highlighting differences in fatigue development between younger and older drivers.

Chapter 6 provides a comprehensive discussion of all the experiments conducted, aiming to integrate complementary theories into a conceptual framework that helps explain the results across the various studies.

Chapter 2 An Investigation of Executive Function in Younger and Older Adults – Implications for Takeover Request Functionality in Autonomous Vehicles

2.1 Introduction

Driving is a dynamic task requiring cognitive, perceptual, and motor functionalities. Declines in any of these functionalities could lead to reduced driving performance and potentially hazardous outcomes (Mäntylä et al., 2009). Inattention, distraction, cognitive workload, etc., therefore, have been the focus of much driver behaviour research. Ageing is a common cause of declines in cognitive, perceptual, and motor capabilities, potentially impacting driving performance (Grady, 2012; Hedden & Gabrieli, 2004; Verhaeghen & Basak, 2005; Verhaeghen & Cerella, 2002). Current studies suggest that markers of ageing start to become noticeable around the age of 40 and become pronounced by 60 years (Lehallier et al., 2019; Shen et al., 2024).

Older adults (60 years and above) face unique challenges in driving, particularly at intersections requiring rapid decision-making and navigation (Preusser et al., 1998). Age-related declines in cognitive processing and motor abilities contribute to slower reaction times, which negatively impact driving performance and increase their vulnerability in complex scenarios (Mäntylä et al., 2007). Studies show that drivers aged 85 and older have the highest age-adjusted fatal intersection crash rate, nearly double that of those aged 70–74, while drivers aged 70–74 are 20–21% more likely to be involved in fatal intersection crashes compared to those aged 20–24 (Lombardi et al., 2017). Although older drivers tend to adopt safer driving behaviours, such as avoiding speeding or risky manoeuvring, their higher injury susceptibility introduces a "frailty bias" in crash statistics, potentially overstating their involvement in severe crashes (Hakamies-Blomqvist, 1998; Cicchino, 2015; Hakamies-Blomqvist et al., 2004).

These issues have been explored in greater detail in the first chapter, where the discussion provides a comprehensive analysis of the issues and challenges of older drivers and their performance in Level 3 autonomous vehicles. Based on the review of the existing literature, it was recommended that further investigation be conducted into the specific cognitive functions necessary for responding to a takeover request in a safe, timely, and effective manner. This includes examining key executive functions, which play a crucial role in ensuring a successful transition from automated to manual control.

2.1.1 Problem Statement

The future integration of Level 3 AVs into the traffic infrastructure necessitates a deeper understanding of the cognitive processes underlying TORs to investigate their impact on driver behaviour and safety. As TORs demand quick shifts in attention, situational awareness, and motor responses, the role of executive functions (EF) becomes crucial in determining successful transitions from NDRTs to active vehicle control.

2.1.2 Study Objectives

1. Develop a conceptual framework linking EF to TORs.
2. Experimentally investigate the impact of ageing on EF performance in TOR-like tasks.

2.2 Conceptual Framework: Linking EF to Takeover Requests

2.2.1 Executive Functions and Takeover Request

Executive function is a collection of cognitive skills essential for the planning, initiation, sequencing, and monitoring of complex goal-directed behaviours, such as driving – a multifaceted activity demanding cognitive, perceptual, and motor coordination (A. Miyake et al., 2000; Salthouse, 2005). While defining EF within the context of driving is challenging due to its intricate nature, studies often identify these key components as updating, shifting, and inhibition (Mäntylä et al., 2009; Akira Miyake et al., 2000; Salthouse, 2005). Studies have shown that there exists a link between driving performance and EF. Mäntylä et al. (2009) found that the updating component of EF was the primary predictor of driving performance, with poorer performance observed in participants with weaker updating abilities. Another study found that impairments in the updating and mental flexibility (shifting) components of EF significantly predicted driving safety, explaining 53% of the variance in performance (Ranchet et al., 2013). Other studies have corroborated these findings, linking deficits in updating and shifting abilities to reduced driving performance (Adrian et al., 2011).

This study conceptualizes TORs through the lens of EF, as illustrated in Figure 2.1. In the context of AVs, the importance of task switching (or shifting) in takeover performance was first highlighted by Zeeb et al. (2017). Later, Shi et al. (2022) utilized task-switching theories to predict and explain takeover performance. These studies focused on using one component of EF, i.e. shifting in this instance, to understand takeover performance. Another significant contribution was made by Scharfe and Russwinkel (2019), who proposed an ACT-R cognitive model that illustrated the cognitive dynamics of automated driving (Scharfe & Russwinkel, 2019). While, Peng et al. (2022) specifically examined how EF, influenced by age, affects takeover performance. They found a significant correlation between EF and takeover performance (Peng et al., 2022).

Drawing inspiration from these works, particularly the contributions of Scharfe and Russwinkel (2019) and Peng et al. (2022), this study presents a conceptual framework linking EF and TORs. A successful response to a TOR involves re-orienting or “shifting” attention to driving, visual scanning of the area and “updating” their understanding of the environment, and “inhibiting” irrelevant stimuli, followed by executing an appropriate motor action, such as steering or braking. Figure 2.1 decomposes the TOR process into three interrelated executive function (EF) mechanisms, represented by distinct colour schemes: grey (NDRT engagement), blue (Shifting), and orange (Updating).

1. **Driver “shifts” their attention from NDRT (Shifting):** While the driver is engaged in an NDRT (processing and encoding information), they remain alert for the initiation of a TOR. This NDRT could be visual (e.g., video watching) or non-visual (e.g., audio listening). Although this engagement may lead to a decline in situational awareness, the system presumes the driver remains mentally available to respond to a TOR and thus may not be completely out-of-loop. Upon initiation of a TOR, attention must shift away from the NDRT and toward the driving environment. This involves an initial disengagement from the task, followed by a visual reorientation toward the road scene, scanning of relevant areas, and localized attentional shifts toward potential hazards.

The perception and processing of the TOR itself is governed by both bottom-up and top-down attentional mechanisms. Bottom-up processing, which is stimulus-driven, plays a critical role in the initial detection of the TOR alert – whether visual (e.g., flashing dashboard signal), auditory (e.g., alarm sound), or a multimodal combination. These salient cues are designed to capture attention reflexively, even when cognitive resources are engaged elsewhere. Conversely, top-down processing supports the interpretation and integration of the TOR within the broader driving context. It enables the driver to suppress the ongoing NDRT, assess environmental demands, and initiate appropriate planning and response behaviours. Thus, effective takeover performance relies on a dynamic interplay: bottom-up processes facilitate rapid TOR perception, while top-down control enables strategic disengagement and re-engagement with the primary driving task.

2. **Driver “updates” (encodes) relevant information (Updating):** Drivers must “update” their mental representation of the environment by encoding relevant information. After reorienting their attention, they visually scan the driving scene, making targeted fixations to gather and process critical data. This process builds situational awareness, enabling drivers to assess the urgency and complexity of the situation effectively.

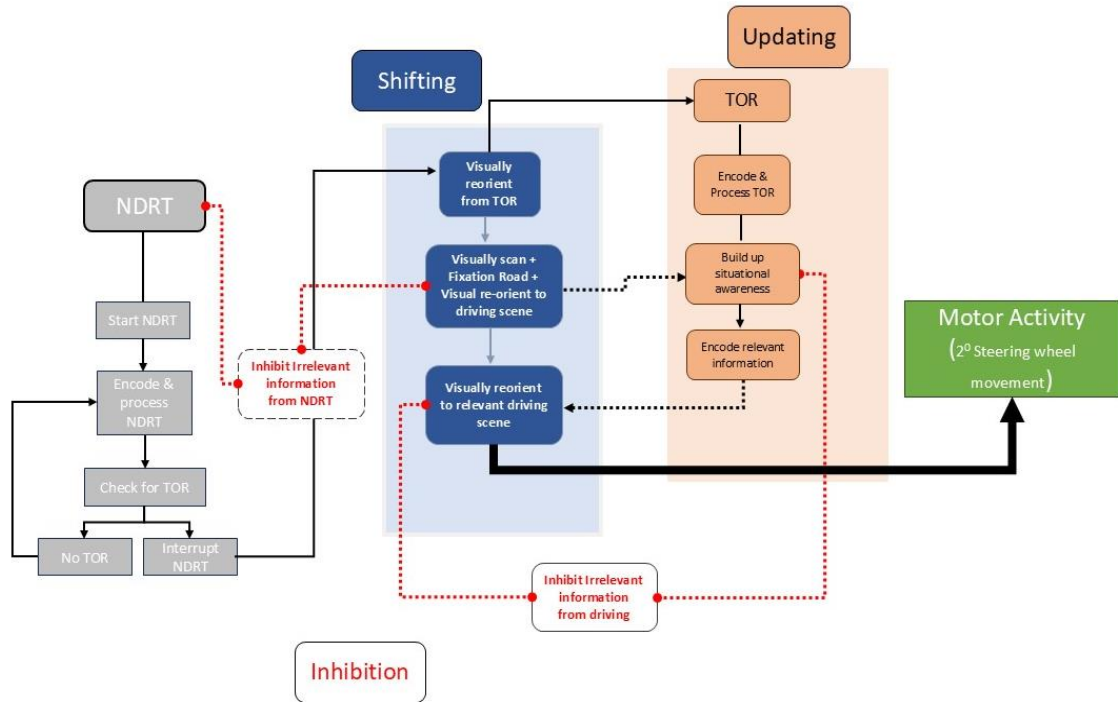


Figure 2. 1. Depicting different elements of executive functioning and attention involved during a TOR scenario.

3. **Inhibiting irrelevant information (Inhibition):** Drivers must “inhibit” irrelevant information both from the NDRT and the driving environment. Initially, distractions from the NDRT must be suppressed to fully engage with the TOR. During the visual scanning process, drivers continue to inhibit irrelevant stimuli from the driving environment to maintain focus on task-relevant cues.

These processes culminate in motor activity, where the driver executes the physical actions necessary to regain control of the vehicle (e.g., steering or braking). Together, shifting, updating, and inhibition form the cognitive foundation for effective takeover performance, emphasizing the crucial role of EF in mitigating risks during TOR scenarios.

By integrating these findings, this framework highlights the interplay between EF components and the behavioural responses required for TORs, offering insights into how cognitive flexibility and attention control impact takeover success in Level 3 autonomous vehicles.

2.3 Experimental Design: Investigating Age-Related Effects on EF

As individuals age, they often experience some declines in cognitive abilities across various domains, including EF (Verissimo et al., 2022). While it is well-established that age-related cognitive decline negatively impacts performance on cognitive tasks, less attention has been paid to differences in the specific subcomponents of EF (updating, shifting, and inhibition) that may be affected by ageing, particularly the degree to which these subdomains differ with age (Maldonado et al., 2020). Understanding these differences is critical, particularly in contexts like TORs, where EF components directly influence task success based on literature and our conceptual framework.

Despite numerous studies on EF, challenges persist in defining and measuring its subdomains. Some researchers employ a single task to explore multiple EF aspects using latent modelling (Packwood et al., 2011), while others conceptualize EF as a collection of distinct subdomains such as updating, shifting, and inhibition (Friedman & Miyake, 2017; Maldonado et al., 2020; A. Miyake et al., 2000). Moreover, even in studies focused on broad subdomains of EF – such as inhibition (Rey-Mermet et al., 2018; Schuch, 2016), shifting (Ashendorf & McCaffrey, 2008; Verhaeghen, 2011), updating (Van der Linden et al., 1994; Zeintl & Kliegel, 2010; Zuber et al., 2019) – older adults (OA) typically exhibited performance deficits compared to younger adults (YA). However, some of these differences may be attributed to general age-related slowing rather than specific deficits in attention or EFs (Salthouse, 2010).

Importantly, while executive function subdomains have been studied individually, there is limited research directly comparing whether they decline uniformly with age or show differential vulnerability. Recent work by Maldonado et al. (2020) highlights the need for such comparisons, showing that some EF subcomponents may be more susceptible to age-related decline than others. Addressing this gap requires a quantitative, domain-specific analysis to determine which subcomponents are most affected and to clarify their functional implications in older adults. By investigating these differences, this chapter aims to provide insights into the cognitive mechanisms underlying TOR performance and the extent to which ageing impacts EF subcomponents.

2.3.1 Description of Takeover Tasks

Scanning the outside environment, also known as visual search, is a crucial skill in driving and TOR scenarios. To incorporate this skill into our study, a visual search task was chosen as the foundational component (Treisman & Gelade, 1980). By combining it with one of the three identified EF components, we derived three distinct tasks, as described below. In this study, the approach was to identify fundamental cognitive elements that would be essential for performing a TOR and to develop tasks based on those components, to investigate age differences.

The visual search task used in this experiment for all four task conditions, showed a target which the participants were supposed to identify. In each trial, participants were required to identify a designated target. This target remained the same for a set number of trials before changing. For example, as shown in Figure 2(a), the initial target might be a yellow car; after several trials, the target would change – such as to a red car – which participants would then need to identify in subsequent trials.

2.3.1.1 *Updating Task*

The ability to update working memory representations is a critical component of EF and is important for many everyday tasks, including driving (Wilhelm et al., 2013). This cognitive process involves continuously refreshing and manipulating information in working memory to adapt to new situations or tasks (Mäntylä et al., 2009). In this study, updating was assessed using a visual search paradigm with adaptations incorporated from continuous memory updating tasks (e.g., n-back). At the beginning of each trial block, participants were shown a “target” image, which they were later required to identify amidst a set of distractor images (see Figure 2.2e). To engage updating processes, the target image was periodically replaced every few trials, requiring participants to actively refresh and maintain the new target in working memory (see Figure 2.2a).

2.3.1.2 *Shifting Task*

The Shifting Task measured the ability to transition attention between different tasks or mental sets. It is a key component of EF and is important for many real-world tasks, including driving. The shifting task employed here drew on classic task-switching paradigms (e.g., Miyake et al., 2000) and dual-task designs used in driving research (Merat et al., 2012). Participants alternated between two tasks: solving simple verbal math problems (simulating a conversational NDRT) and performing a visual search to identify a target image among distractors. Prior to each visual search trial, a green directional arrow cue (presented for 700 milliseconds) was used to spatially orient attention toward the location of the target image, based on the Posner cueing paradigm (Posner, 1980) (see Figure 2.2b). These cues served a dual function:

1. Reorienting participants' spatial attention toward the target location.
2. Priming participants for false cues was introduced in the Inhibition Task (description below).

This dual-task design replicated real-world driving scenarios where drivers must divide attention between a primary driving task and secondary activities, such as conversations (Merat et al., 2012). It also highlighted the mental demands of switching between cognitive tasks (see Figure 2.2b).

2.3.1.3 Inhibition Task

The Inhibition Task focused on participants' ability to suppress irrelevant or misleading information. This task was evaluated using a paradigm inspired by negative priming frameworks and cue-conflict designs. It drew inspiration from the concept of negative priming, where the participants who were primed to rely on particular information (orienting cues in this case) usually showed slower responses/or more errors when the nature of that information was changed (e.g., the target becoming a distractor) (Chao et al., 2022). Tipper et al. (2001), while critiquing the contemporary inhibition models of negative priming, rejected the notion that negative priming does not involve an inhibitory mechanism (Chao et al., 2022; Tipper, 2001). This approach used here is consistent with that adopted by Tipper et al., (2001).

Here, participants were again required to identify a target image amongst the distractor images. Similar to previous tasks, the target would change after each trial block. As the participants were

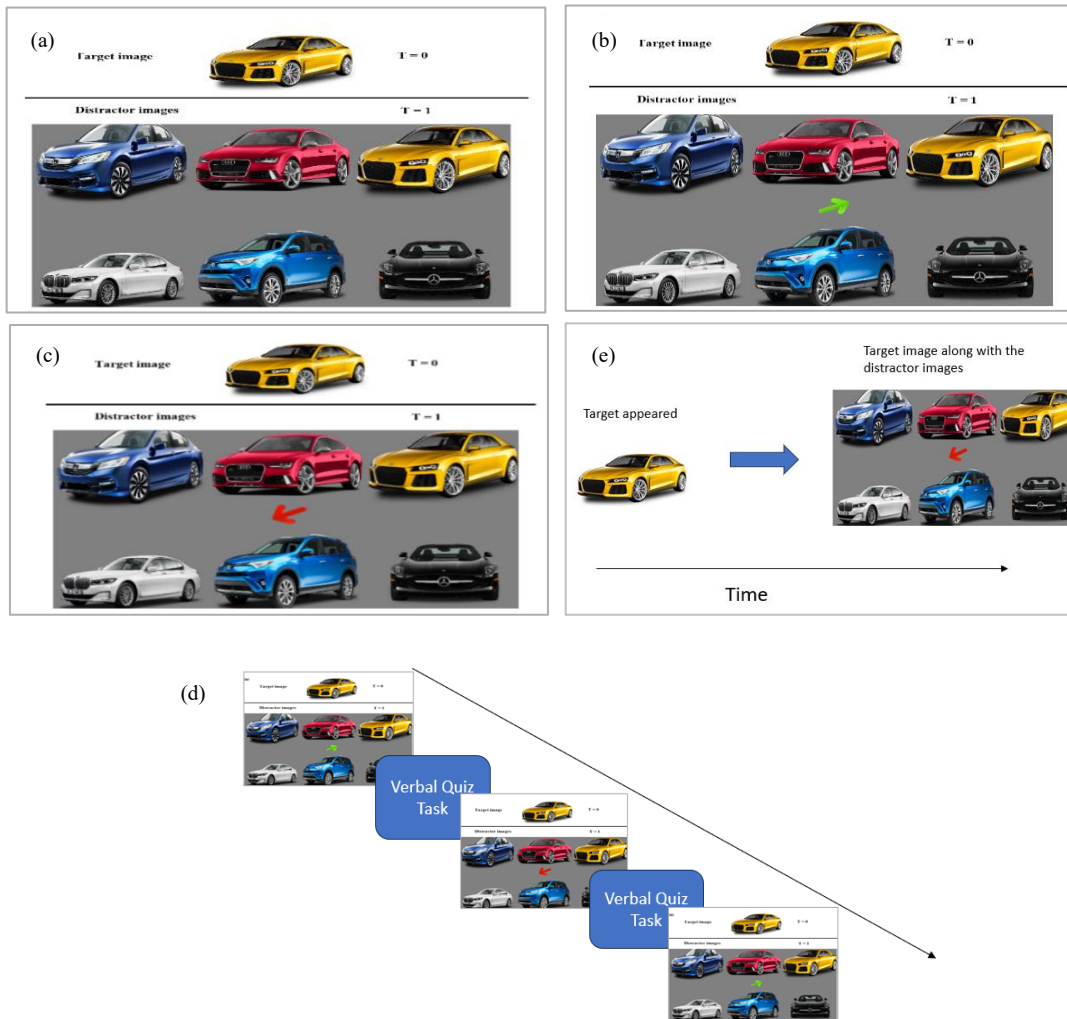


Figure 2. 2. A snapshot depicting (a) Updating Task, (b) Shifting Task, (c) Inhibition Task during the experimental run of these tasks, (d) Combined Task, and (e) describing the appearance of “Target” image along with “Distractor” images.

primed to rely on information from orienting cues (which provided correct spatial information about the target) in the shifting task, incorrect information provided using similar arrows would have to be inhibited. Red arrows were used as false cues (which provide incorrect information about the spatial location of the target) along with legitimate orienting cues (green arrows), similar to the Posner cueing task, see Figure 2.2(c). The proportion of false cues to the orienting cues was in the ratio of 3:10.

2.3.1.4 *Combined Task*

Since a TOR engages multiple EF components simultaneously, as shown in Figure 2.1, this task aimed to integrate these elements. The Combined task in this study includes the following elements:

1. Verbal task
2. Visual search task
3. Orienting and False cues

These elements represent different subcomponents of EF: updating through target changes after each trial block, shifting between the verbal and visual search tasks, and inhibition through the false cues. In a real takeover scenario, various cues (such as auditory, visual, or tactile) alert drivers to disengage from NDRT and re-engage with driving. In our study, we assumed a visual cue, though research is ongoing to determine which modality is most effective.

In the Combined task, a sequential dual-task paradigm was used. Participants first completed a verbal quiz, followed by a visual search task. Before the visual search, cues were provided that could either be orienting cues or false cues. This serial task sequence mimicked the temporal flow of a takeover situation, where drivers must disengage from a cognitive distraction, process incoming cues, and visually re-engage with the environment under potential conflict or uncertainty (figure 2.2d).

The assumed task difficulty hierarchy in this study – that is, Shifting as more demanding than Inhibition, and Inhibition as more demanding than Updating – was justified on two complementary grounds: first, based on the theoretical complexity of the EF components themselves; and second, based on the design characteristics of the tasks used to operationalize these components.

From a theoretical perspective, Shifting is widely regarded as the most cognitively demanding EF component. It involves disengaging from one mental set and reconfiguring for another, often under time constraints and in the presence of interference. This requires considerable top-down control, coordination, and attentional flexibility. Neuroimaging studies have shown that task-switching engages broader and more distributed brain networks, particularly within the

dorsolateral prefrontal cortex and parietal regions, reflecting the extensive executive resources required (Dove et al., 2000; Monsell, 2003). In contrast, Inhibition – while also effortful – typically involves suppressing dominant or misleading responses within a stable task set. Although age-related declines in inhibitory control are well documented (Rey-Mermet et al., 2018), the overall control demands remain lower than those associated with set-shifting. Updating, by comparison, is generally the least demanding of the three EF processes. It involves the continuous replacement and monitoring of information in working memory without the need for switching task sets, and is considered less susceptible to age-related decline (Miyake et al., 2000; Zelazo et al., 2008).

This theoretical hierarchy was reflected in the task design used in the current study. The Updating Task was designed as a simplified n-back-like paradigm embedded within a visual search task. It involved repetitive stimulus monitoring and minimal interference, with no requirement for switching or alternating cognitive sets. The Inhibition Task introduced greater complexity by requiring participants to suppress spatially incongruent cues, though the task itself remained a single-stream design with consistent stimulus-response rules. In contrast, the Shifting and Combined Task was deliberately structured as a serial dual-task paradigm in which participants had to alternate between a visual search task and a verbal decision task. This imposed frequent task-rule reconfiguration, dual-modality integration, and increased working memory and attentional demands. Such designs have been shown to elicit higher response times, error rates, and frontoparietal activation in past research (Braver et al., 2003; Worringer et al., 2019).

2.4 Research Questions

The following are the research questions that were investigated:

RQ 1. To what extent does the complexity of EF-based tasks (updating, shifting, and inhibition) influence participants' ability to detect (in terms of speed and accuracy) a preselected stimulus?

H1. As the complexity of EF-based tasks increases (from updating to inhibition to shifting and combined tasks), participants' ability to detect a preselected stimulus will decline, as reflected by longer reaction times and higher error rates.

RQ 2. How do age-related differences in EF subcomponents – updating, shifting, and inhibition – impact reaction time and accuracy in tasks designed to simulate the cognitive demands of TORs in Level 3 autonomous vehicles?

H2. Older participants will exhibit slower reaction times and higher error rates across all EF tasks compared to younger participants, reflecting age-related declines in cognitive processing speed and attentional control. Age-related performance differences will be most pronounced in the

combined task, as it requires the integration of multiple EF subcomponents, exacerbating cognitive limitations in older adults.

2.5 Method

2.5.1 Design

The experiment consisted of two parts: a primary study (Experiment One) and a follow-up study (Experiment Two) to control for potential fatigue effects and validate the manipulation of task difficulty.

Experiment one employed a 4×2 mixed-factor design, with Task Condition (Updating, Shifting, Inhibition, and Combined) as the within-subjects factor and Age Group (Younger, Older) as the between-subjects factor. Reaction time and accuracy served as the dependent variables.

Experiment two focused exclusively on the within-subjects variable (Task Condition) and involved only younger participants. The sequence of task conditions was altered to rule out fatigue and sequence effects as potential confounds.

Evidence of fatigue as a potential confound in Experiment One came from the increasing trend in reaction times (RT) across tasks, progressing from Updating to Shifting, Inhibition, and Combined. While this pattern aligns with increasing task difficulty, the sequential presentation of tasks raised the possibility that fatigue, rather than complexity, contributed to slower responses. Fatigue is known to impair cognitive performance, potentially leading to increased RTs.

To address this, Experiment Two interchanged the presentation order of the Inhibition and Updating tasks. The assumption was that if the RT still followed the presentation order then fatigue would likely be the primary cause.

2.5.2 Participants

In experiment one, a sample of thirty-six healthy participants was recruited using different advertising strategies such as posters, emails, contacts through local organizations, etc. The sample statistics calculation was done on G-power software using F -tests with Power, $\beta = 0.8$ and Type I error, $\alpha = 0.05$, and effect size, $f = 0.5$. These participants belonged to two age groups (18 in each group): Younger adults (18 to 35 years old) and Older adults (60 and above years old). Most younger ($M = 23.11$, $SD = 5.12$) adults were students at the University of Dublin, Trinity College, while the older adults ($M = 66.9$, $SD = 4.32$) were local community-dwelling individuals recruited through personal contacts and various other advertising, see Table 2.1.

In experiment two, twenty younger adults ($M = 19.28$, $SD = 1.21$) were recruited through advertisement at the University of Dublin, Trinity College. This experiment two was conducted to mitigate the confounds from fatigue or presentation order. If the same participants from

experiment 1 had been involved in experiment two, their familiarity with the tasks could have introduced learning effects, where participants become more efficient or quicker at completing the tasks simply due to repeated exposure. By using a new set of participants, we ensured that each individual was approaching the tasks with no prior experience, allowing us to isolate the effects of task order more effectively. Additionally, recruiting a new set of young participants for experiment two allowed us to replicate the findings of experiment one while ensuring the results were not specific to a particular sample. This has improved the generalizability of these findings.

All participants were provided with a detailed information sheet and informed consent form before the experiment, in compliance with the guidelines of the Trinity College Dublin School of Psychology Research Ethics Committee with approval id - SPREC012022-15. Participation in this study was entirely voluntary and participants received either a one4all¹ voucher worth 20 Euros or two research credits in the case of students. The exclusion criteria for participation included: a history of seizures or epilepsy, currently taking psychoactive medications; or having any kind of nonremovable metal near their head. No participant was excluded based on these exclusion criteria.

2.5.3 Materials

The experiment was developed using PsychoPy software (version 2021.2.3) in a window-based environment. The experimental set-up included a Dell XPS laptop, a Bush 24-inch monitor screen (with dimensions H33.4, W55.3, D6.3cm) with a resolution of 1366 x 768 pixels an external keyboard, and a mouse used by the participants to respond.

2.5.4 Measures

2.5.4.1 *Reaction Time*

Reaction time was defined as the time elapsed between the onset of a stimulus and clicking on the mouse to identify the target.

2.5.4.2 *Accuracy*

Accuracy has been defined as several trials in which the target is correctly identified. For analysis purposes, we have used error rate, defined as the total number of trials minus the total number of correct trials.

2.5.4.3 *Perceived Task Difficulty*

To comprehend the effect of task conditions on performance, a measure of subjective task difficulty was employed in the second experiment only. Task difficulty refers to the participant's subjective score based on their assessment of task complexity (Nawaz et al., 2020). After each

¹ One4all is a multi-store gift card.

task, participants were asked to score the task difficulty level and the mental demands they experienced during that particular task. Mental demand was explained to the participants as the number of attentional resources required for them to complete a task. The following are the questions that were asked of the participants after the completion of each task-

- a. How difficult do you think this task was on a scale of 1 to 10, 1 being very easy and 10 being very hard?
- b. How mentally demanding do you think this task was on a scale of 1 to 10, 1 being very easy and 10 being very hard?

The participants responded to these questions verbally which were recorded using pen and paper. The perceived task difficulty was defined as the average value of the two subjective measures that were asked of the participants.

2.5.5 Experimental Procedure

Upon entering, participants were briefed on the experimental setup, including the visual search task, target changes after each trial block, and the dual-task nature of the shifting and combined tasks. After verbally confirming their understanding, informed consent was obtained. This procedure was consistent for both experiments. ~~The accuracy was determined by counting the total number of correctly identified target images.~~

2.5.5.1 Experiment One

Once the informed consent was taken, in experiment one, participants were given four practice trial blocks, which were based on each task condition (i.e. Updating block, Shifting block, etc.), to make them familiar with the setup and experimental task. Each practice trial block consisted of five trials where the participants were required to identify the target image. The images used in the practice trials were distinct from those which were used in the experiment. Given the simplicity of the experimental tasks, the majority of participants consistently achieved accuracy

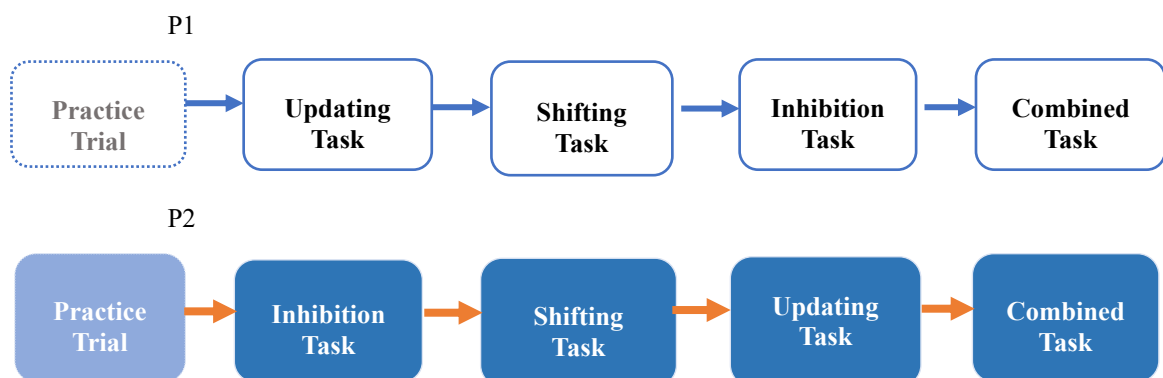


Figure 2. 3. Showing the sequence of the tasks presented in experiment one (above one). (P2) the sequence of tasks presented to the participants in experiment two.

rates surpassing 95% during the practice phase. Consequently, the actual experiment commenced after the mutual satisfaction of both the experimenter and the participant with their performance in the practice trials. After the practice trial, there were four experimental task conditions: Updating, Shifting, Inhibition, and Combined task respectively. These tasks were presented in the order of their inferred increasing difficulty, see Figure 2.3 (P1), with a short break of 2-4 minutes after each task however, even though Shifting was assumed to be more difficult than Inhibition, it was presented earlier. This was done to ensure that participants would be primed to rely on the information provided by the cues (as orienting cues) during the Shifting task. This would make it likely that participants came to rely on the cued information making it necessary for the participants to inhibit the cues when the primes provided the wrong spatial information. The Combined task was a complex arrangement of the first three tasks and thus followed after the individual tasks. Completion of all task conditions marked the end of Experiment One for each participant.

Each task condition had in total of 48 trials and each task had six trial blocks, with each trial block comprising eight trials with the same stimulus. This stimulus changed after each trial block. The participants were required to identify the target image from a set of distractors across eight trials after which the target was changed. This general structure was maintained across all task conditions, each exhibiting varying completion times. Participants accomplished the shortest completion time in the Updating block ($M = 2.35$ minutes, $SD = 0.63$), while the longest completion time was noted in the Combined task ($M = 15.37$ minutes, $SD = 0.73$). The total testing time was approx. 36 minutes, where updating and inhibition tasks were 4 to 6 minutes long while shifting and combined tasks were roughly 10-12 minutes long.

2.5.5.2 *Experiment Two*

Once participants signed the informed consent form, they completed four practice trial blocks, each corresponding to a different task condition (i.e., Updating block, Shifting block, etc.), to familiarize them with the setup and tasks. After both the experimenter and the participant were satisfied with the participant's performance in the practice trials, the experiment began. Experiment two was conducted as a follow-up study to address the issue of fatigue. The order in which the four tasks were presented to the participants was changed to Inhibition, Shifting, Updating, and then Combined tasks, see Figure 3 (P2). Once all task conditions were completed, the experiment for that participant concluded.

Also, to fully comprehend and explain the results from both of these experiments, a subjective Perceived Task Difficulty score was used to quantify task difficulty and mental demands experienced by the participants. To aid participants in accurately rating task difficulty a baseline task was provided. This baseline task only had a visual search task and did not have any of the

distinctive elements of the four different experimental tasks (such as changing of target after each trial, verbal task, or cues). This ensured that the difficulty level of this baseline task was very low. Participants were specifically told that they had to report the difficulty/mental demands of the other tasks in comparison to the difficulty/mental demands of the baseline task.

The tasks were completed in the order shown in Figure 2.2 (P2), with a short break of 2-4 minutes after each task condition. Additionally, in Experiment Two, task demand was subjectively measured at the end of each block. The basic structure of these tasks stayed the same as in Experiment One. The total time it took participants to complete the experiment was also similar to around 36 minutes.

2.6 Results

2.6.1 Experiment One

A 4×2 mixed factor Analysis of Variance (ANOVA) was conducted using SPSS (v. 28.0.1.1) on each average reaction time and accuracy with Task Condition as a within-subject factor and Age, as the between-subjects factor. The effects from the factors were considered significant for a $p < 0.05$. The descriptive statistics and ANOVA tables of the results have been provided in Appendix A. Participants' demographic information is reported in Table 2.1.

Table 2. 1. Participants' demographic and educational background for Experiment One.

Age group	Gender	Mean Age	Undergraduate	Postgraduate	Others	Total number
Old	Male	68.30	6	4	1	11
	Female	64.28	4	2	1	7
Young	Male	22.68	6	7	2	15
	Female	25.33	0	3	0	3

2.6.1.1 Reaction Time

The reaction time data were first analysed for the presence of outliers. Individual reaction times of less than 150 milliseconds were considered outliers and were removed (Thompson et al., 1992). The upper limit for reaction time to be considered an outlier was the maximum time duration for each trial, which was 4.8 seconds. The data removed on this basis were less than 0.02% of the total.

The data were expected to violate the assumption of the equality of variance (Levene's test of variance) because of the increase in inter-individual variability with age (Hale et al., 1988). The results from Levene's test of equality of variance were found to be significant, however, there is

strong evidence to support that ANOVA is robust enough to overcome such violations (Budescu, 1982). It was found that both Age and Task Condition had significant effects ($p < 0.001$) on the performance in EF-based tasks, see Figure 2.4. Since no interaction between Age and Task Condition was observed, the main effects were interrogated.

2.6.1.1.1 The main effect of Task Condition

Mauchly's Test of Sphericity was significant, therefore, F -statistics were reported using Greenhouse-Geisser corrections. The effect of the task was significant, $F(2.107, 71.627) = 65.721$,

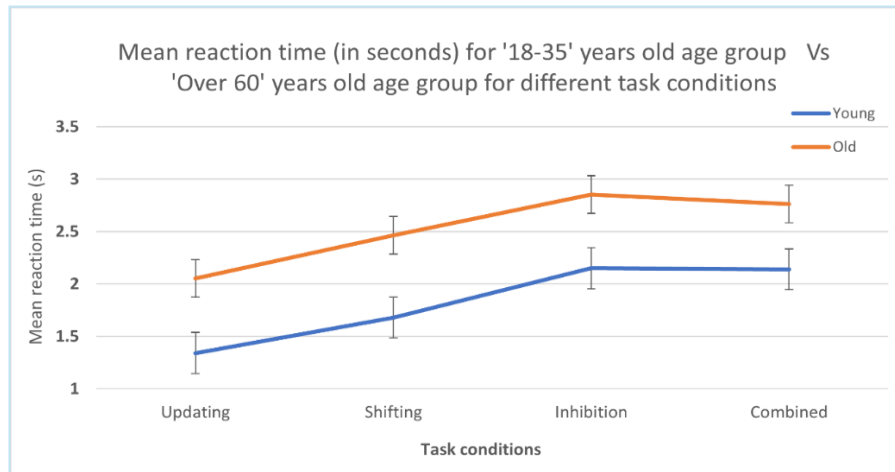


Figure 2.4. Mean reaction time (in seconds) for 'Over 60' years old and '18-35' years old adults across each tasks i.e. Updating, Shifting, Inhibition, and combined tasks. The error bars here represent 95 % confidence interval.

$p < 0.001$, $\eta_p^2 = 0.659$. Pairwise tests with Bonferroni corrections indicated that comparisons between Updating, Shifting, Inhibition, and Combined tasks were significant ($p < 0.001$) for each pair except for Inhibition-Combined.

2.6.1.1.2 The main effect of Age

The effect of Age on reaction time corroborated previous findings and showed that there was a significant effect of age, $F(1, 34) = 23.871$, $p < 0.001$, $\eta_p^2 = 0.412$. It can be observed from Figure 2.4, that the older adults' reaction times were higher than younger adults across all tasks.

2.6.1.2 Accuracy

In this study, the participants were required to find the target image amongst the distractor images by clicking on it. Clicking on any other image was recorded as an incorrect response. The number of correct and incorrect responses was counted for each task and each age group, see Figure 2.5. Participants who made six or more incorrect responses (on average, one error per block) were considered outliers and were removed from the accuracy data analysis. Five participants, from both groups combined (two from the young and three from the old age group), were subsequently removed.

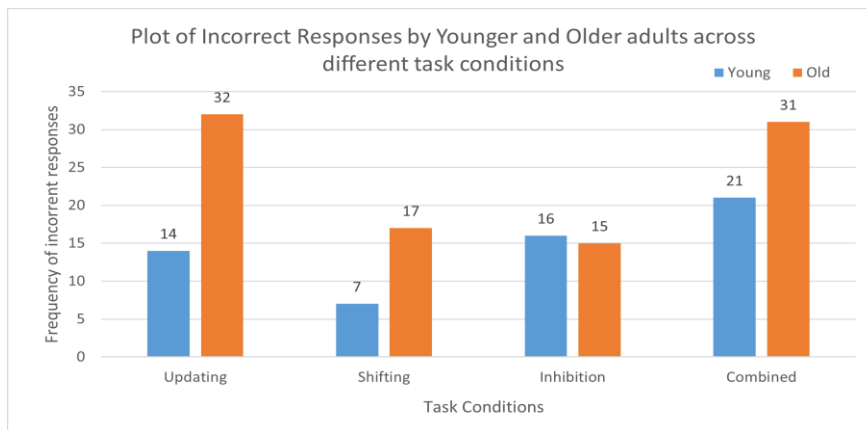


Figure 2. 5. Plot of the number of incorrect responses for each task across both age groups.

2.6.1.2.1 The main effect of Task condition

Since Mauchly's Test of Sphericity was not significant, F -statistics are reported with sphericity assumed. The effect of task condition on the frequency of incorrect responses was found to be significant $F(3, 102) = 3.438, p = 0.02, \eta_p^2 = 0.092$. Pairwise tests with Bonferroni correction found that only the comparison of the Shifting and Combined condition was found to be significant ($p = 0.026$). Note the interaction between Task condition and Age was non-significant $F(3, 102) = 3.438, p = 0.297, \eta_p^2 = 0.035$.

2.6.1.2.2 The main effect of Age

The effect of Age on the frequency of incorrect responses was also found to be significant, $F(1, 34) = 6.078, p = 0.019, \eta_p^2 = 0.152$, with older adults producing more mistakes (total = 95) than younger adults (total = 58).

2.6.2 Experiment Two

2.6.2.1 Reaction Time

Experiment two only had one group of participants, so, the effect of the task condition (with 4 levels) was considered using a within-subjects ANOVA. Outlier removal was conducted in the

same way as for experiment one. The effects from the factors were considered significant for a $p < 0.05$.

Since Mauchly's Test of Sphericity was significant, F -statistics were reported using Greenhouse-Geisser corrections. The effect of the task condition on the reaction time was significant, $F(2.342, 44.495) = 123.637$, $p < 0.001$, $\eta_p^2 = 0.867$. This effect can be observed visually in Figure 2.6, where the average reaction time had been plotted for each task condition. In experiment two, the higher effect size is likely due to the greater homogeneity in the sample concerning age ($M = 19.28$, $SD = 1.21$). There was a significant effect of the task condition on the performance, the

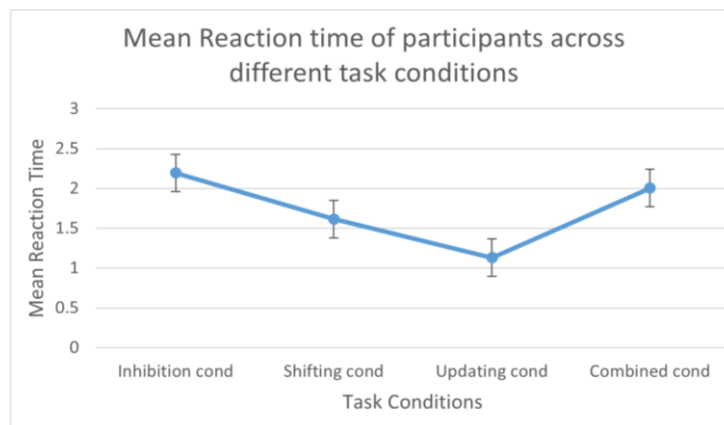


Figure 2. 6. Mean Reaction time for each task as observed in Experiment two.

pairwise comparisons between tasks with Bonferroni correction were considered as a post hoc test to pinpoint the origin of the significant results. The results of task condition on reaction time performance from the post hoc test were similar to experiment one ($F(2.107, 71.627) = 65.721$, $p < 0.001$, $\eta_p^2 = 0.659$), where all task pairs showed a significant effect ($p < 0.001$) on performance except for the Inhibition-Combined task pair ($p = 0.057$).

We also conducted a factorial ANOVA analysis, treating participants from two experiments as a between-subject variable, to examine the impact of the order of presentation of conditions (indicative of fatigue) on reaction time performance. The results revealed a non-significant effect of the order of presentation (or fatigue) on reaction time performance, $F(1, 36) = 2.146$, $p = 0.152$, $\eta_p^2 = 0.056$. Conversely, a significant effect of Task Condition on reaction time emerged, $F(2.057, 74.046) = 143.964$, $p < 0.001$, $\eta_p^2 = 0.800$, with Greenhouse-Geisser corrections applied to the F -statistics. Importantly, these findings were in line with our expectations. Additionally, figure 2.7 presents a comprehensive view of the results by illustrating the z-scored normalized reaction time, cumulative incorrect responses, and task difficulty levels on a single graph, facilitating a holistic understanding of the findings.

2.6.2.2 Accuracy

The results are reported for Huynh-Feldt correction as Mauchly's test of sphericity was significant ($p < 0.05$). The analysis showed a statistically significant effect of task condition on the frequency of incorrect responses $F(2.383, 45.286) = 7.768, p < 0.001, \eta_p^2 = 0.29$. The results showed a

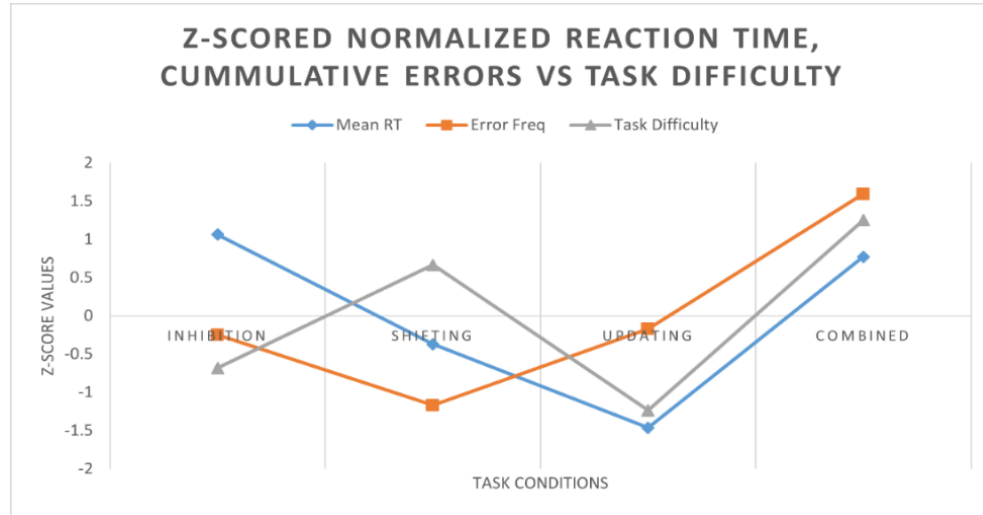


Figure 2. 7. Plot of Z-scored normalized data for Reaction Time, Cumulative Errors Vs Task Difficulty.

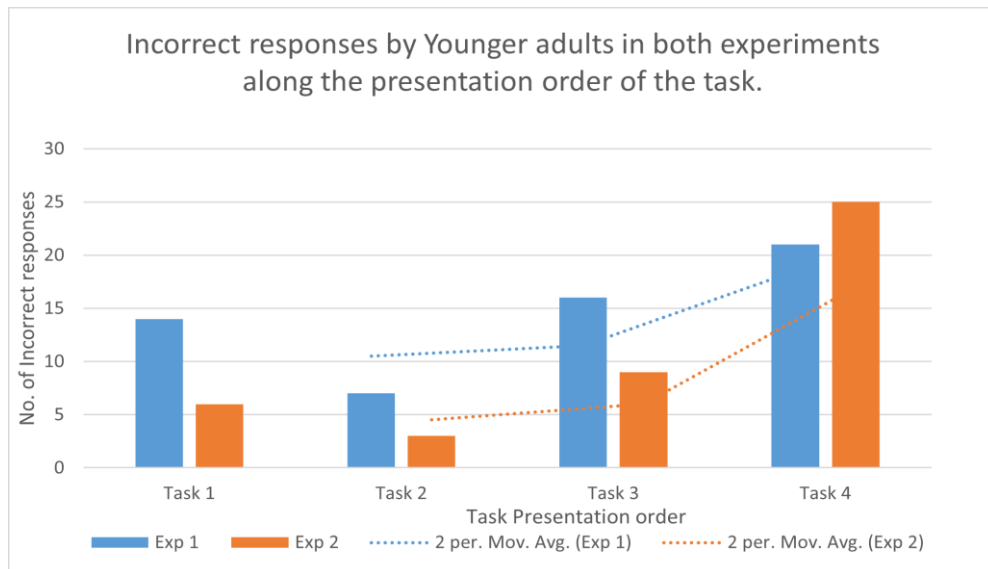


Figure 2. 8. Frequency plot of incorrect responses by participants for each task in the Experiment two and by only younger participants from the Experiment One. The tasks are named in the order of their presentation.

significant effect of task condition on the frequency of incorrect responses, pairwise comparison tests with Bonferroni correction have been conducted as a post hoc test. The results showed that only two condition pairs were significant i.e. Shifting-Combined condition ($p = 0.021$) and Inhibition and Combined condition pair ($p = 0.003$) showed a statistically significant difference in results, while every other task pair failed to show statistical significance, see Figure 2.8.

2.6.2.3 Task Difficulty

Repeated measure ANOVA was conducted on the perceived Task Difficulty scores of the participants. The results were statistically significant for task condition $F(3, 54) = 88.638$, $p < 0.001$, $\eta_p^2 = 0.831$. A pairwise comparison was done with Bonferroni correction, which showed that the results were significant for all task condition pairs. In Figure 2.9, mean perceived task difficulty scores were plotted against task conditions.

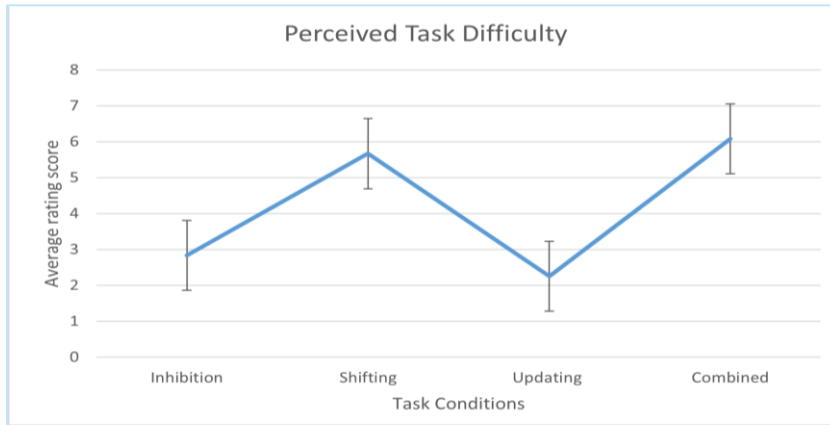


Figure 2. 9. Mean Perceived Task difficulty scores across task conditions in the presentation order.

2.7 Discussion

2.7.1 Theoretical Implications – Conceptual Framework

This study contributed a novel conceptual framework that situated EF at the core of effective TOR performance in Level 3 autonomous vehicles. The framework emphasized three interrelated EF components – shifting, updating, and inhibition – each playing a critical role in the cognitive demands of a TOR scenario. Peng et al. (2022) found a significant correlation between EF and lateral driving stability, but they did not propose a structured framework linking EF to TOR performance. This study addressed that gap by conceptualizing how each EF component contributes to a successful takeover. By aligning with prior research on EF and driving (Mäntylä et al., 2009; Scharfe & Russwinkel, 2019; Peng et al., 2022), this framework provides a comprehensive lens for understanding the cognitive underpinnings of engagement and disengagement processes in the lead-up to TORs, offering theoretical and practical insights into designing systems that support driver cognition during automated driving.

While these EF components may affect TOP in automated driving, this study did not directly assess their impact during actual takeover scenarios. As such, caution is warranted in extrapolating the effects observed in these tasks to TORs.

2.7.2 Experimental Implications

Hypothesis 1: Task performance in the series of tasks would follow the same pattern as their respective task difficulty, which, was assumed to be as follows: Updating would be the easiest one then inhibition, shifting, and finally combined.

2.7.2.1 *Reaction Time Analysis (Experiment One)*

Hypothesis one was partially validated. While reaction time did vary across the conditions the sequencing diverged from our predictions. Both age groups exhibited an increase in reaction time when traversing from Updating to Shifting, and then to Inhibition, except for the Combined task, where it plateaued. Contrary to expectations, the reaction time for Shifting proved faster than for Inhibition. Conceptually, Shifting and Combined tasks, given their dual-task paradigm, were expected to be more demanding, incurring higher cognitive loads and consequently having slower reaction times. Therefore, it is interesting to note that the performance in the Shifting task was faster than in the Inhibition task. These results were in line with the findings of Idowu and Szameitat (2023), where they found inhibition to have the highest rates of decline (thus poorer performance), followed by shifting and updating.

Similarly, the Combined task – presumed to be the most difficult task – should have the slowest reaction time and yet the reaction time plateaued. The reaction time of the Combined task was similar to the inhibition task, despite the Combined task being a dual task while the inhibition task was a single task. This unexpected pattern might be attributed to variations in mental workload, especially concerning demands on working memory. The Updating task elicited the quickest responses due to its minimal cognitive demand. The Shifting task, grounded in a dual-task (DT) paradigm (comprising both verbal quiz and visual search elements), placed greater cognitive demands on participants. Yet, the introduction of orienting cues may have mitigated these demands by reducing the working memory load. Comparable reaction times for Inhibition and Combined tasks may reflect similar cognitive loads imposed by false cues in both tasks. This suggests that false cues in the Inhibition task might have imposed a cognitive load akin to managing multiple tasks along with false cues in the Combined task. Notably, participants' ability to switch focus between verbal and visual elements in the Combined task without delay suggests the absence of a "switching cost" relative to the Inhibition task. Furthermore, the presence of orienting cues in the Combined task may have reduced the cognitive load, and when combined with learning effects, this may have allowed participants to manage task switching more effectively, thereby reducing the anticipated "switching cost" and thus improving their RT.

No studies to date have attempted to merge multiple subcomponents of EF into a single, unified task, to the best of the authors' knowledge. As a result, it is challenging to directly compare findings from the 'combined task' – which integrates various EF components – with those from

other studies that have examined EF as a collection of distinct, independent tasks. Most existing research tends to explore EF subdomains (such as inhibition, shifting, and updating) separately, treating each as a standalone cognitive process. Without a framework that combines these elements into a single task, understanding how different subcomponents interact or overlap in real-time cognitive performance remains limited. This gap in the literature makes it difficult to assess whether the combined task reveals unique insights about EF that might differ from findings based solely on individual subcomponent analysis. Furthermore, without comparable studies, it is unclear how merging EF subcomponents in a single task impacts cognitive performance, or whether this approach more accurately reflects the interconnected nature of these functions in everyday scenarios.

There could be an alternative explanation for this increase in reaction times across the task – slower reaction times on later blocks could have been due to fatigue instead of increased task difficulty. To address this confound the experiment was repeated with just younger participants and the task sequence was changed such that the order of the updating and inhibition were switched, see Figure 2.3. Despite changes in task order, reaction times exhibited the same pattern observed in Experiment One: the fastest responses were recorded for the Updating task, followed by the Shifting task, while the Inhibition and Combined tasks showed comparable reaction times. A detailed discussion of these findings is provided in the following sections.

2.7.2.2 *Accuracy Analysis (Experiment One)*

Accuracy was analysed using a mixed-factor ANOVA. Initially, higher error rates indicated a learning curve during the Updating task, with a marked reduction in errors during Shifting. One plausible reason for this reduction is the introduction of orienting cues. These cues directed participants' attention spatially toward the intended target, thereby facilitating more accurate responses and minimizing errors. However, a contrasting trend was noted in subsequent tasks: error rates climbed notably from the Inhibition to the Combined task, as visualized in Figure 2.5. This increment in errors during the Inhibition and Combined conditions could be ascribed to a couple of factors. First, in the Inhibition task, false cues might introduce additional cognitive load, effectively diverting attention and leading to errors. The DT nature of the Combined task coupled with false cues compounded cognitive demands, further increasing errors.

The accuracy or incorrect responses were affected more by age than task conditions as evidenced by the effect sizes for task condition 0.092 and 0.152 for age respectively. The observed effect size for reaction time (RT) was found to be 0.659. In contrast, the effect size for accuracy stood at a modest .092, characterizing it as a relatively small effect. This indicates that the complexity of the task is better revealed by reaction time, with accuracy giving us additional but limited

information. Managing these parallel demands could strain cognitive resources, resulting in a higher likelihood of errors (Barra et al., 2006; Lyell et al., 2018).

Hypothesis 2: Age difference in performance will be observed, where older adults would have slower reaction time but would have a comparable error rate to the younger adults as the task difficulty increases.

2.7.2.3 *Ageing Effects on Reaction Time and Accuracy:*

It was observed that the reaction times of older adults were consistently slower than their younger counterparts across all tasks. Such findings align with existing research, where older adults often exhibit age-related declines in areas like reaction times, working memory, mental rotation, etc. (Kausler, 1991; Salthouse, 1996). Researchers have suggested this slower reaction time could be because of older adults' reduced information-processing capacities (Salthouse, 1996). These findings are also in line with more recent studies that investigated the comparison of ageing effects within the subdomains of EFs (Idowu & Szameitat, 2023; Maldonado et al., 2020). The results for the older adults' reaction time performance for different EFs were in line with our expectations given our understanding of task difficulty, except for the shifting-inhibition and inhibition-combined pair. The effect size for the reaction time data for Task Condition was 0.659 and for Age was 0.412. The fact that the effect size was smaller for age indicates that Task Condition was a bigger source of variation than Age. In addition, it is important to note that since there was no interaction among the data, the impact of Age was effectively consistent across all four Task Conditions, which means, the advantage of being younger is consistent irrespective of task.

The results suggested that the inhibition component of EF experiences greater age-related decline compared to other EF components. This aligns with the findings of Idowu et al. (2023) and is supported by the inhibition deficit theory (Hasher & Zacks, 1988), which posits that older adults are more vulnerable to distracting interference due to reduced attentional control. This reduction in inhibitory resources with age impairs the ability to filter out irrelevant information, making it harder to retrieve task-relevant data. A recent fMRI study by Saylik et al. (2022) showed that while all EF components activated overlapping brain areas, particularly in the prefrontal cortex (PFC), each also engaged distinct regions. This may explain the different rates of decline observed in specific EF components. As this system allocates cognitive resources, it highlights the potential overlap in EF processing, especially in working memory, inhibition, shifting, and updating. The diversity of EFs, involving additional brain regions like the parietal cortex, may contribute to age-related cognitive decline, advancing our understanding of cognitive ageing.

This comparison of reaction time between the EF tasks has not been extensively studied. Therefore, these findings may offer a new perspective to study EF and ageing and further our understanding of cognitive ageing.

Accuracy trends were similar but nuanced, see Figure 2.5. Particularly, there was no interaction between Age and Task Conditions concerning participants' error rate (incorrect responses). It's worth clarifying that the error rate in Figure 2.6 represents the absolute error rate, which is the cumulative number of errors made by one age group across all trials. In comparison, the relative error rate – often cited in research – amounts to an average of 0.0329 errors per trial per participant for the older group. This number is very small, leading us to report the aggregate errors. For younger participants, this average is 0.0188. Therefore, older individuals exhibited a slightly elevated error rate when identifying target images. The reason for such a low error rate is that the tasks in themselves were very simple and thus would not warrant a higher error rate.

Similar to younger participants, older adults also showed a decline in performance when traversing from the updating to the shifting condition; however, the error rate was significantly higher in older adults compared to their younger counterparts. Others have reported similar results, where the older adults' accuracy performance was poorer in comparison to younger adults (Bolton & Staines, 2012; Kenemans et al., 1995). Unlike younger participants, whose error rate increased from shifting to inhibition, older adults' errors plateaued at the inhibition stage. In the combined condition, older adults exhibited a much higher error rate than younger participants. These findings contradicted those of McLaughlin and Murtha (2010), who observed that older adults tend to make greater use of contextual information, such as cues, to enhance their performance (McLaughlin & Murtha, 2010). In contrast, the results from current study indicate that orienting cues, particularly when combined with false cues, actually worsened the accuracy of older adults. This counterintuitive effect may be attributed to the elevated cognitive demands of inhibiting misleading cues, while simultaneously shifting between the quiz and search tasks and accurately identifying the target. The combined load on inhibition, task-switching, and selective attention likely overwhelmed older adults' executive resources, resulting in reduced accuracy. The effect size for accuracy data for the main effect of Task Conditions was 0.092 and for Age was 0.152. These effect sizes suggest that Age, despite with low effect size, had a higher impact on the accuracy performance than the Task Condition.

Throughout the study, it became evident that older adults, potentially due to cognitive declines and slower information processing, demonstrated both lengthier reaction times and a higher average error rate than their younger counterparts. Yet, it's encouraging to see that the older group's performance improved, reflected in diminishing error rates, as they familiarized themselves with the experiment. However, their performance declined during the combined task, which had a greater task load than any previous single task. Fatigue might also have played a role in this decrease.

Overall, this study's findings resonated with existing literature, reinforcing the notion of slowed reaction times and cognitive declines in older adults (Grady, 2012; Hedden & Gabrieli, 2004;

Verhaeghen & Basak, 2005; Verhaeghen & Cerella, 2002). However, our results add nuance by showing that age-related deficits were not uniform across all tasks. Older adults struggled most with tasks requiring inhibition and complex task-switching, particularly under dual-task or combined conditions, suggesting a selective vulnerability in these EF domains.

These findings have important implications for the design of human-machine interfaces (HMI) in Level 3 autonomous vehicles. TOR systems could be optimized by allowing more time for decision-making, reducing misleading cues, and providing adaptive support tailored to cognitive load. Such design considerations could mitigate the disproportionate burden placed on older drivers during high-demand takeover situations.

Experiment Two: Investigating confounds of Fatigue

Task Difficulty:

To address potential fatigue confounds, Experiment Two altered the task order and introduced subjective Perceived Task Difficulty scores. ANOVA results confirmed significant differences among tasks, with Updating perceived as easiest, followed by Inhibition, Shifting, and Combined. This subjective order diverged from reaction time data, where Shifting was faster than Inhibition. The divergence may result from limited conscious processing of false cues in the Inhibition task, influencing perceived difficulty.

We analysed these values using repeated measures ANOVA which showed significant differences between task conditions. Moreover, pairwise comparisons revealed the difference was significant amongst all task pairs' perceived difficulty scores. These differences showed that there are discernible variations in how participants subjectively interpret and assess the cognitive demands associated with different task conditions. Task difficulty score identified Updating to be the easiest, followed by Inhibition, Shifting and finally Combined. This order doesn't align with reaction time data, where, Shifting had the second-fastest reaction time after Updating, which should have been Inhibition as per Task Difficulty. The divergence between reaction time and task difficulty at the inhibition could be explained by of lack of/limited conscious processing in the inhibition of false cues (i.e. red arrows). This lack of /limited conscious processing would then influence a participant's perceived experience of task difficulty (Van Gaal, Ridderinkhof, Scholte, & Lamme, 2010). This interpretation aligns with findings by van Gaal et al. (2010), who demonstrated that inhibitory control processes – particularly those triggered by No-Go signals – can be activated unconsciously, engaging prefrontal networks without entering conscious awareness. In our study, participants – especially older adults – showed slower reaction times and reduced accuracy during inhibition trials. However, their subjective difficulty ratings did not reflect a proportionally high perceived workload. This discrepancy suggests that, while the inhibitory demands were cognitively taxing, participants may not have been fully aware of the

effort (in performance deterioration) required to suppress false cues, particularly when inhibition was triggered automatically. Yet, the overall subjective appreciation of the task difficulty was consistent based on a high effect size $\eta_p^2 = 0.831$. Nonetheless, it could be argued that the Perceived workload score may be relevant in terms of understanding the perception of cognitive demands but it fails to capture the dynamic fluctuations of the cognitive load which is being reflected in the reaction time. Some of the limitations of the Perceived Task Difficulty are the subjective nature of the self-reports that could be influenced by biases, the cognitive and mental state of the participants, social desirability bias, etc. (Martens et al., 2016). A review article reported a low to medium correlation between subjective measures and direct measures (Prince et al., 2008). Analysing both subjective and more direct metrics such as reaction time and accuracy allows for a more comprehensive understanding of participants' engagement, proficiency, and cognitive processing strategies within each task condition. However, given the lack of a standardized index that could accommodate these different modalities i.e. self-reports, accuracy, and reaction time data, we are inclined to adopt reaction time as a more valid measure of task performance.

Reaction Time:

As previously discussed, experiment two altered the order of the Updating and Inhibition tasks, as illustrated in Figure 2.3. ANOVA analysis results indicated that the effect of task conditions is significant even after altering the presentation order of the task conditions. Despite altering task order, reaction times followed the same pattern as in Experiment One: fastest for Updating, followed by Shifting, with similar times for Inhibition and Combined tasks. This challenged the thesis that fatigue might be affecting the increase in reaction time in moving from one task to another in the order of their presentation in experiment one. While fatigue might have had some impact, the dominant influence on reaction time shifts was the task condition. This assertion is supported by the reaction times in experiment two, coupled with a non-significant effect of the experimental group (i.e. young adults from experiment one compared to those from experiment two) in the factorial ANOVA. It is important to note that the effect sizes from both experiments (when only considering young adults in experiment one) were comparable (η_p^2 for experiment one – 0.747, η_p^2 for experiment two – 0.867), indicating reaction time to be a more robust indicator of task performance. Given the lack of impact of task sequence on reaction time, it seems unlikely that fatigue impacted these results.

Accuracy:

Error rates showed no significant differences between task orders, confirming that fatigue was not a primary factor influencing accuracy. Despite the Shifting task being perceived as more difficult, error rates appeared to be decreasing from Inhibition to Shifting (non-significant

difference), with a notable spike in the Combined task. This surge suggests an increase in the cognitive load in the Combined task and the dip in error rates at the Shifting task would suggest a decrement in the cognitive load. Note the distribution of incorrect responses across participants indicates that the deviation in error rates is not attributable to a few outliers but rather represents a systematic trend.

If error rates were a product of factors other than individual task loads, changing the presentation order would have affected them. However, post-hoc analyses across Experiment One and Experiment Two demonstrated no significant differences in error rates when task orders (e.g., switching the order of Updating and Inhibition) were altered, further supporting this conclusion.

Despite the lack of statistical significance in post-hoc comparisons, discernible trends in error rates were evident across conditions. Both experiments showed a general tendency for error rates to increase with task order progression, except for a dip in the Shifting task, as illustrated in Figures 2.5 and 2.7. This pattern suggests that mental fatigue may have played a role in influencing error rates, alongside cognitive load. Notably, these factors impacted participants' target recognition abilities, even though reaction times (Figure 2.8) remained unaffected. These findings align with Faber et al. (2012), who proposed that mental fatigue increases the processing of irrelevant information, thereby reducing accuracy.

When considering results from both experiments, a nuanced interplay between cognitive load and mental fatigue emerges as a factor influencing performance, which warrants further investigation. Additionally, the findings highlighted inconsistencies across different performance metrics in reflecting cognitive demands under various task conditions as depicted in Figure 2.7.

As discussed above, when careful examination of Figure 2.8 was undertaken, a notable observation emerged: there exists a remarkable consistency in younger participants' data. Specifically, Task 2 exhibits the lowest performance metric, succeeded sequentially by Task 1, Task 3, and Task 4 (the numbering here followed the presentation order). Such a trend implied that the measure of accuracy may serve as a more indicative metric of learning progression and mental fatigue levels, rather than solely reflecting cognitive load. Thus accuracy as a metric is unveiling more about the presentation order of the task and age than the cognitive load.

From both Experiment One and Experiment Two, Reaction time was found to be the most sensitive metric to assess the changes in cognitive load as mentioned above. Notably, it remained unaffected by mental fatigue; this fact is highlighted by similar reaction time corresponding to different task conditions in experiment one and experiment two as well as by comparable effect size of reaction time between experiment one (when only taking younger participants) and experiment two.

2.8 Conclusion

This study sought to investigate the relationship between EF subcomponents – updating, shifting, and inhibition – and TOR performance in autonomous vehicles, with a focus on age-related differences. The results provided important insights into how ageing influences reaction time and error rates across tasks with varying cognitive demands.

As hypothesized, older adults exhibited slower reaction times across all tasks compared to younger adults, consistent with previous research on age-related cognitive declines (Salthouse, 1996; Kausler, 1991). While task condition had a stronger impact on reaction times than age, the persistent slowing in older participants underscores the universal cognitive slowing associated with ageing. The most significant decline was observed in inhibitory control, supporting the inhibition deficit theory (Hasher & Zacks, 1988). This is further corroborated by studies that have demonstrated distinct, age-sensitive brain regions involved in EF processing, particularly in the prefrontal cortex (Saylik et al., 2022).

In terms of accuracy, both age groups followed a similar performance pattern across tasks. However, older adults showed higher error rates, particularly in complex conditions like the Combined task, indicating difficulty in managing diverse cognitive demands simultaneously. The presence of false cues appeared to exacerbate these age-related declines, deviating from previous findings that older adults utilize contextual cues to enhance performance (McLaughlin & Murtha, 2010).

The results demonstrated that Age had a negative impact on the tasks conducted within this research. It also highlighted how the complexity of EF tasks can exacerbate these declines. If the tasks were a valid representation of the subcomponents of a TOR then this may indicate potential difficulties for older drivers engaging in a TOR. Further research will be required to confirm the extent to which these tasks represent components of a real-life TOR.

It is crucial to note that these results do not suggest that older adults cannot handle TOR scenarios or would have bad takeover performance because they perform slower than younger adults and have made more errors (in general) than younger adults in this experimental task. It has been widely noted that results from simple experimental tasks don't necessarily translate to complex real-life situations such as driving, where the person's experience, muscle memory, cognitive abilities, etc., all contribute to their driving performance (Holleman et al., 2020). Ensuring an adequate response window for TORs can help accommodate older drivers' needs without compromising safety. Therefore, it would be reductive to draw any strong conclusions about older adults driving from this study at this stage.

In summary, our findings contribute to the growing body of literature on cognitive ageing, offering new perspectives on EF task performance and ageing, particularly within the domain of

autonomous vehicle research. Future research should explore strategies to mitigate the cognitive demands on older drivers during takeover scenarios, potentially incorporating adaptive cues and assistive technologies to ensure safer driving experiences for all age groups.

While this study primarily focused on age-related decline in EF, the EF-based task framework also has meaningful implications for younger, inexperienced drivers. Despite possessing intact or even superior cognitive resources compared to older adults, younger drivers may lack the driving-specific schemas and situational awareness needed to effectively deploy those executive functions during takeover scenarios.

In particular, inexperienced drivers may struggle with integrating task-switching (shifting), real-time information updates (updating), and suppression of irrelevant or misleading cues (inhibition) when faced with high-pressure transitions from automation to manual control. These limitations are not due to EF deficits per se, but to the absence of procedural knowledge and contextual judgment that comes with driving experience.

Therefore, EF-based measures could help identify not only age-related vulnerabilities but also competence gaps in younger drivers, especially in novel or ambiguous takeover situations. The findings support the development of driver training programs or adaptive systems that scaffold EF engagement for younger drivers by simulating high-load TOR events and providing feedback on attention allocation, cue prioritization, and response selection..

2.9 Implications for L3 AVs

While this study did not directly investigate TORs in L3 AVs, the findings offer preliminary insights. False cues were shown to impair reaction times, emphasizing the importance of minimizing distractions and providing clear, unambiguous information during TORs. Age-related declines in inhibitory control suggest the need for adaptive interfaces or extended transition times for older drivers to compensate for reduced attentional control.

Reaction times were more sensitive to task complexity than accuracy, highlighting the need to streamline information and reduce cognitive demand during TORs. The study's focus on EF subcomponents also provides a potential framework for assessing driver readiness. However, further research is required to develop tasks that better replicate the dynamic, multimodal demands of real-world TOR scenarios.

2.10 Limitations of the Current Study

This study has several limitations that warrant consideration. First, the participant pool was small and lacked diversity across age groups, limiting the generalizability of the findings. A larger, more representative sample could provide deeper insights into age-related differences in EF and TOR

performance. Second, the experimental tasks may not have been sufficiently challenging to elicit a wide range of errors, particularly in participants with higher cognitive abilities. Increasing task complexity could yield more robust data on error patterns and reaction times. Additionally, this study did not explore dynamic fluctuations in cognitive load or the impact of mental fatigue, both of which are critical during real-world tasks. Finally, the findings are not validated with driving simulations or real-world driving scenarios, limiting their ecological validity. While the controlled laboratory setting allows for isolating cognitive processes, it does not fully capture the complexity of real-world driving.

Therefore, the next chapter report on an experiment conducted using a high-fidelity driving simulator to examine the impact of time pressure – specifically, the time budget available during a takeover request – on takeover performance. The study aims to understand how variations in the allotted time budget influence drivers' ability to regain control of the vehicle safely and efficiently. Additionally, a key focus is placed on investigating the role of ageing in takeover performance, exploring whether age-related differences in cognitive and motor abilities contribute to variations in drivers' responses under different time budgets.

Chapter 3 Age Differences in Takeover Performance Across Varying Time Budgets in Level 3 Autonomous Vehicles

3.1 Introduction

The ability of drivers to successfully regain control after a Takeover Request (TOR) depends on several factors, including the Time Budget (TB) – the interval provided for takeover – and demographical factors, such as age (Gold et al., 2013; Zeeb et al., 2017). Understanding how these factors exert their influence and interact with each other is crucial for designing autonomous vehicles (AV) systems that ensure safe and efficient human automation interactions.

3.1.1 Effect of Age on Takeover Performance (TOP)

We concluded in the previous chapter that older drivers may take a longer response time to a stimulus due to their slower information processing capabilities. Moreover, age-related differences in driving and cognitive abilities are well-documented, with older adults typically exhibiting slower reaction times, reduced cognitive flexibility, and diminished sensory processing, all of which can hinder their ability to swiftly and effectively take over control from an AV. Given these declines with age, it is reasonable to assume that older drivers would find TOR scenarios more challenging, potentially leading to poorer TOP.

Current literature, however, presents a conflicting perspective regarding the impact of age on TOP. Several studies have found that older drivers exhibit slower reaction takeover time (RTOT), greater variability in steering control, and overall poorer takeover performance than younger drivers, particularly under time-constrained conditions (Li et al., 2018, 2019; S. Li et al., 2021; Peng et al., 2022; Wu et al., 2020). These findings suggest that age-related cognitive decline makes older drivers less efficient in responding to TORs, especially under time pressure.

Conversely, other studies have found no significant difference in TOP between younger and older drivers (Clark & Feng, 2017; Körber et al., 2016; Miller et al., 2016). Research by Körber et al. (2016) investigated the influence of age on takeover performance when drivers were engaged in a questioning task presented by a hands-free phone. The study found age had no effect on takeover time (TOT), but older subjects braked more frequently and harder and maintained greater time to collision (TTC). The authors interpreted these findings as evidence of more cautious takeover behaviour among older adults, likely attributable to their greater driving experience. In addition, research by Clark and Feng (2017) investigated age differences in the preferences for non-driving-related tasks as well as takeover performance. They found that older and younger drivers preferred to engage in different NDRTs during automated driving. Younger subjects were more likely to use electronic devices, while older subjects were more likely to talk to other people. In some cases,

research has even suggested that older drivers outperform younger ones, likely due to their greater situational awareness and more cautious approach to driving (Favarò et al., 2019). These studies suggest that older drivers, despite cognitive declines, can compensate with experience, cautious driving behaviour, and strategic decision-making, allowing them to perform takeovers as efficiently as younger drivers.

3.1.2 Effect of TB on TOP

Research in the field of automated vehicles indicates that TB significantly affects driving performance during a TOR. Research has shown that an excessively low TB can result in poorer takeover performance and an increased likelihood of collisions (Jamson et al., 2008; Q. Li et al., 2021; Wan & Wu, 2018). Meanwhile, an excessively long TB can result in a lack of acceptance of autonomous driving vehicles, because the driver might consider the TOR as a false alarm or a nuisance (Jin et al., 2021). In a study by Mok et al. (2015), drivers were presented with TBs of 2, 5, and 8 seconds to navigate a construction zone. They found only a few participants managed to navigate when presented with a 2-second TB, while most of the drivers, with lead times of 5 and 8 seconds exhibited slower but more successful and deliberate manoeuvring (Mok et al., 2015). Additionally, most drivers in both the 5-second and 8-second TB conditions managed to safely navigate road hazards. Further research has examined how different time budgets affect driving performance. For instance, when comparing 4-second and 8-second TBs, longer TBs were associated with better driving outcomes, such as successful lane changes and fewer collisions (Dogan et al., 2021). Various studies have tested different TBs based on the vehicle automation technology and the specific scenarios, 5 seconds vs. 7 seconds (Gold & Bengler, 2014), and 4 seconds vs. 8 seconds (Dogan et al., 2021; Yousfi et al., 2018). A common finding across these studies is that shorter TBs are generally associated with shorter response takeover time, higher workload, and increased safety risks (Zhang et al 2019; Gold et al; Mok et al., 2015). In contrast, longer TBs tend to facilitate smoother transitions but may still pose challenges depending on the cognitive and perceptual state of the driver (Marberger et al., 2018).

While much of the research in automated vehicles suggests that a TB of 5 seconds and longer would be optimal for safety, there is still inconclusive evidence regarding what constitutes a “high” or a “low” time budget (Gold et al., 2013; Shahini et al., 2023; Wan & Wu, 2018). Studies have also focused on 3, 4, 5 seconds, etc. TBs, demonstrating that while drivers generally perform well at longer TBs, lower ones, present significant challenges (Mok et al., 2015; Wan & Wu, 2018). However, research by Gold et al. (2013) showed that drivers can respond in as little as 2.06 seconds. Additionally, a meta-analysis conducted by Zhang and colleagues found that takeover times could range up to 20 seconds, with an average of 2.72 seconds, which further supports the notion that drivers may be capable of handling lower TBs (Zhang et al., 2019). Therefore, determining an optimal range for lead time/time budget is crucial, balancing the trade-off between

early and late warnings to ensure safety, effective response, and driver trust. These findings highlight the importance of systematically investigating how different TB durations impact takeover performance in realistic driving scenarios.

This inconsistency in the literature raises important questions about the underlying factors influencing TOP in older drivers. One potential explanation for these discrepancies is the role of the TB provided for takeover. While shorter TBs may exacerbate age-related deficits, extended TBs could allow older drivers to compensate for slower cognitive processing, potentially reducing performance gaps (Soares (Körber et al., 2016; Soares et al., 2021). However, prolonged TBs may also lead to automation complacency, delaying responses in both younger and older drivers. This highlights the need for a deeper investigation into the interaction between age and TB in TOP.

Despite extensive research on age-related driving performance and TB effects, little is known about how these factors interact in Level 3 AVs. Specifically, it remains unclear:

1. How does age influence takeover performance in Level 3 AVs?
2. How does the available time budget affect takeover performance across different age groups?

By addressing these questions, this study aimed to resolve inconsistencies in the literature and provide insights for designing age-inclusive AV systems that optimize TOR settings and TB allocation to enhance safety and usability for all drivers. This study was conducted as two experiments, where the first experiment investigated the above-mentioned research questions while the second experiment was conducted in a more advanced driving simulator with different demographics to repeat the findings from the first experiment. Both experiments had a Time Budget as an independent variable which had two levels: 2.5 seconds (Low TB) and 5 seconds (High TB), while experiment one also investigated age as an independent variable.

3.2 Method

3.2.1 Experiment Design

Two experiments were conducted as a part of this chapter and both utilized a within-subjects design to examine the effects of Time Budget on driver performance and workload. The dependent variables encompassed behavioural measures obtained from the driving simulator (shown in table), as well as subjective assessments of workload and fatigue, measured using the 5-point NASA-TLX scale. Each participant was presented with four task conditions and each task condition had two takeover events. These takeover events would be based on the combination of time budgets – high TB or low TB, see Table 3.3. All calculations for the metrics reported in Table 3.3 were performed from the onset of the TOR notification up to the moment immediately preceding the successful completion of the lane change by the participant. All calculations for the metrics

reported in Table 3.3 were performed from the onset of the TOR notification up to the moment immediately preceding the successful completion of the lane change by the participant.

Table 3. 1. Takeover Performance Metrics

Dependent Variables	Levels	Definition
Steering Wheel Angle (SWA) - Degrees (°)	at 2°	The standard deviation of the degree of rotation applied to the steering wheel by the driver after the TOR has been initiated.
Steering Wheel Angle Speed (SWS) - (°/s)	at 2°	The rate of change in the steering wheel angle (SWA) over time, indicating how quickly the driver adjusts steering after the takeover request (TOR).
Reaction Takeover Time - Seconds (RTOT) (s)	at 2°	the time interval between the system's takeover request and the driver's initiation of control actions, i.e. steering.
Standard Deviation of Lane Position (SDLNP) - Meters (m)	at 2°	Standard deviation measures the extent to which the vehicle deviates from the center of its lane during and after the takeover.
Deceleration - (m/s ²)	at 2°	Maximum Deceleration experienced laterally by the vehicle, indicating the smoothness and stability of the takeover maneuver.

3.2.2 Experiment One

3.2.2.1 Participants

A priori sample size estimation was performed using G*Power 3.1.9.4 to determine the minimum required sample size for this study. The calculation was based on a two-tailed alpha of 0.05, a power of 0.8, and an anticipated medium effect size ($f = 0.25$). A medium effect size was selected because prior research in this field has typically reported a medium effect size was considered the smallest effect size of interest (Riesthuis, 2024). Therefore, a total of 32 participants were recruited through snowball sampling and Trinity College Dublin's School of Psychology research system. Participants aged 17–24 years were classified as young drivers (Wu et al., 2020; Department, 2022, November 24), while those aged 60 and above were classified as older drivers (Shen et al., 2024). The study included 17 younger participants and 15 older participants, Table 3.2.

Participant eligibility was determined based on specific inclusion criteria. All participants were required to have a valid driving license, and to have a corrected-to-normal vision, as vision

Table 3. 2. Participants' Demographics for Experiment One.

Participants	Numbers	Mean Age (Years)
Young	17	25.34 ($\pm$ 4.16)
Old	15	66.2 ($\pm$ 4.62)
Total	32	45.76 (SD = 14.76)

impairments could affect their ability to respond to takeover requests accurately. Finally, individuals with major health conditions or a history of severe motion sickness were excluded from the study to mitigate potential confounding effects related to physiological discomfort or pre-existing medical conditions. Participants received two research credits, where eligible, or a €20 One4All voucher. Five participants were excluded from the analysis because they either drove in an incorrect lane or did not finish the Takeover scenario.

3.2.2.2 *Materials*

This study used a fixed-based, right-hand-drive driving simulator consisting of a clutch, brake, accelerator, 7-point gear stick, and force-feedback Fanatec steering wheel, see Figure 3.1. The simulator operated on SCANeR Studio software v2023.1 (AVSimulation, 2022). Three mounted 32-inch monitors were placed in front of the simulator. Speed (km/hr) and current gear were shown on the middle screen.



Figure 3. 1. Driving Simulator along with the mounted Tablet used in this study.

3.2.2.3 *Procedure*

Participants provided informed consent before beginning a practice session to familiarize themselves with the driving simulator. The experimenter also introduced the AV system, experiment format, and TOR to ensure participants understood key terminologies and experimental procedures. During practice, participants drove manually for five minutes in light traffic, following a lead bus while maintaining a safe distance of approximately 3.8 meters. An alert prompted them if they got too close. They were also required to manoeuvre safely around an obstacle in their lane. The session was deemed successful if they consistently maintained a safe distance and navigated the obstruction safely.

The experimental scenario took place on an eight-lane road (four lanes on each side) with mild bends, minimal traffic, and environmental elements like greenery, buildings, etc. (see Figure 3.2(a)). The ego vehicle started in the second lane from the centreline. Participants maintained the practised safe distance and adhered to a manual driving speed limit of 90 km/h, while automation operated at a constant 90 km/h. After 100 seconds of manual driving, automated driving was initiated. The start of automated driving was marked by a visual alert, provided on the top-left of the screen and a minor haptic feedback from the steering wheel. During automated driving, participants were asked to play a game called Subway Princess Runner (IvyMobile, 2023), an endless runner game where participants swipe on the screen to avoid obstacles, presented on a

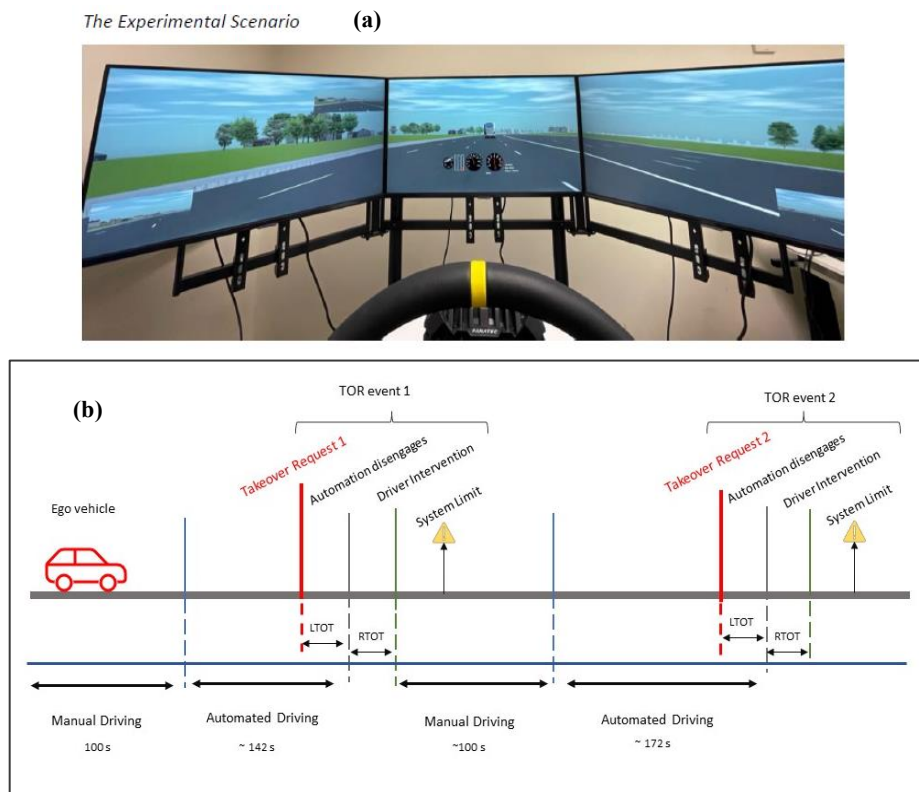


Figure 3. 2. (a) The experimental driving scenario. (b) the schematic diagram of the TOR events in the experiment.

tablet mounted beside the steering wheel. The choice of this particular NDRT was based on insights from a meta-analysis by (Coyne et al., 2023), which concluded that tasks requiring both visual and manual input led to greater physiological arousal compared to those without such demands. This was to understand how varying level of cognitive, perceptual and sensory engagement affect TOP. A sharp beep signalled the TOR, requiring participants to take control and avoid a stationary vehicle in their lane by overtaking safely. Two Time Budgets (2.5s vs. 5s) were tested, with the TOR issued accordingly. This left three seconds until a potential collision, providing sufficient time for a successful takeover, see Figure 3.2(b) (Gold et al., 2013; Portron et al., 2024).

Engagement in automation was indicated by haptic steering feedback and a visual alert. Subjective assessments were conducted using the NASA Task Load Index (NASA-TLX) to evaluate workload, which was administered after each trial, and participants completed a video clip quiz. The entire session lasted approximately one hour, with each trial taking around five minutes, including four minutes of driving and the NASA-TLX assessment. All data were anonymized.

3.2.3 Experiment two

3.2.3.1 *Participants*

A similar sample size calculation was conducted using G*Power for Experiment One. Accordingly, participants were recruited from the FotoNation/Tobii² 24-hour Drowsiness study, in which 29 individuals took part in a 1.5-hour Level 3 automated driving session. Participants were compensated for their time and were required to meet specific health criteria, including no history of major medical conditions or severe motion sickness. Participants were instructed to have more than 7 hours of healthy sleep on the night prior to the experiment to ensure they started the experiment in a well-rested state. They were also advised to abstain from consuming caffeinated beverages on the day of the experiment to avoid the potential confounding effects of stimulants on drowsiness and alertness. Compliance with these instructions was verified verbally at the start of the experimental session. Of the 29 participants, 17 identified as female and 12 as male, with ages ranging from 20 to 56 years ($M = 24.26$, $SD = 7.93$), Table 3.3. Six participants' data were excluded from the analysis because they failed to respond to the takeover request properly as they did not change lanes or drove in an incorrect lane.

Table 3. 3. Participants' demographics for Experiment two.

Participants	Numbers	Mean Age
Male	12	23
Female	17	26.58
Total	29	24.79 ($SD = 7.93$)

² Tobii AB (Previously FotoNation) Ltd an eye tracking company based in Sweden and Ireland.

3.2.3.2 Apparatus

The experiment was conducted using the FotoNation/Tobii in-house, high-fidelity dynamic driving simulator (DS), as shown in Figure 3.3. The simulator features a car interior mounted on an actuated platform equipped with a specialized steering wheel, pedals, and gearstick, all providing force feedback to enhance immersion. Three 49-inch monitors form a continuous, wide-



Figure 3. 3. The Driving Simulator used in Experiment two.

field display that renders driving scenes in front of the simulator. Additionally, a smaller monitor is mounted slightly elevated in front of the driver to display parallel tasks without obstructing the main view. The DS hardware and simulation software are controlled by one computer, while a separate system manages the collection of video, audio, and other data types. The simulator runs on SCANeR Studio software (AVSimulation, 2022).

3.2.3.3 Procedure

Participants arrived at 9 a.m. for a 24-hour drowsiness acquisition session, within which several manual driving sessions were conducted. The experimental session for autonomous driving (L3) began at 9 p.m., lasting 1.5 hours. The manual driving sessions had a different objective and thus had a different methodology, so these are not discussed here. The participants were given sufficient time to rest in between the different sessions. Prior to the L3 session, participants were provided the opportunity to rest for at least 3.5 hours and have their dinner. Before commencing the L3 experiment, participants were briefed on autonomous vehicles, the experimental structure, and their tasks. Participants were also instructed about the experiment procedure, such as maximum speed, lane to drive in, how to respond to TOR, etc. They were then given a practice session to familiarize themselves with the driving simulation environment. The experiment consisted of four driving scenarios.

Each scenario started with approximately 2 minutes of manual driving, followed by the activation of automation for around 2.5 minutes. Two Takeover events were featured in each scenario, with

two types of Time Budgets (Low or High). The timing and sequence of the experiment are shown in Figure 3.2(b). Automation was signalled by haptic feedback on the steering wheel and a visual alert on the top left of the central screen displaying “Automation On.” During automation, the vehicle maintained a speed of 90 km/h. An auditory TOR – a sharp beep – was issued after a set time, instructing the driver to regain control. In all scenarios, the critical situation involved a broken-down car in the vehicle’s lane, requiring the driver to perform an overtaking manoeuvre to avoid a collision. While automation was active, participants were engaged in a similar game as in experiment one called Subway Princess Runner. Two TBs were applied: 2.5 seconds and 5 seconds. . When the TOR was initiated, the automation continued driving for either 2.5 or 5 seconds, allowing the driver to regain situational awareness and mental capacity. After this buffer period, the visual text on the monitor changed, signalling the driver to resume control. A potential collision would occur three seconds after the TOR if no action was taken, see Figure 3.2(b). After successfully overtaking the obstacle, the second Takeover event was activated automatically once the participants’ vehicle had driven for roughly 100 seconds and had crossed over a particular trigger placed on the second lane of the scenario. Following each TOR, participants resumed manual driving until automation re-engaged, indicated by both an "Automation On" text and haptic feedback. In the second TOR event, participants followed a bus that changed lanes to avoid another broken-down vehicle, prompting a second auditory TOR. The modality of the TOR alert varied based on the Time Budget: a low-pitched beep for 5 seconds and a high-pitched beep for 2.5 seconds. Each participant experienced four scenarios, with different combinations of Time Budget settings, as shown in Table 3.4.

Table 3. 4. Describe a Latin Square experiment design where H stands for High TB (5 seconds) and L stands for Low TB (2.5 seconds).

Scenario 1	Scenario 2	Scenario 3	Scenario 4
HH	HL	LH	LL

3.3 Results

The results were analysed using SPSS (v. 28.0.1.1). The method used was mixed factor and one-way Analysis of Variance (ANOVA) using the SPSS for Experiment one and Experiment two respectively. Results are reported separately for Experiment One and Experiment Two. Each TOP metric was calculated from the moment following the TOR notification and subsequent automation disengagement, up to the point when the driver made a 2-degree change in steering wheel angle.

Bonferroni correction was not applied in this study due to its overly conservative nature, which can increase the risk of Type II errors – particularly when outcome measures are correlated, as in

this case (e.g. SWA, SWS, etc.). Instead, uncorrected p -values are reported to preserve sensitivity to potentially meaningful effects. However, we acknowledge that this approach increases the risk of Type I errors, and findings should be interpreted with caution.

3.3.1 Experiment One

A mixed factor ANOVA was conducted on the TOP metrics. The interaction between Time Budget and Age was not statistically significant. Therefore, main effects of TB and Age were analysed and reported separately to examine their independent contributions to takeover performance.

Table 3. 5. Descriptive Statistics and p -value for TOP metrics based on TB as a factor for Experiment One.

TOP Metrics	Mean High TB	Mean Low TB	p -value
Long Deceleration	1.02	1.181	0.464
Lat Deceleration	-0.22	-0.354	0.146
Long SDLnP	235.44	259.57	0.742
Lat SDLnP	134.75	164.70	0.378
SWA	7.38	10.85	0.006*, **
SWS	0.31	0.46	0.005*, **
Vehicle Speed	22.11	23.03	0.433
RTOT	1.623	1.45	0.905

* significant at $p < 0.05$

3.3.1.1 Effect of TB

The effect of TB was analysed across various takeover performance metrics (see Table 3.5) and NASA-TLX, with statistical significance set at $p < 0.05$. The results revealed a significant effect of TB on SWA, $F(1, 25) = 8.964, p = 0.006, \eta^2 = 0.264$, indicating a substantial effect size, Figure 3.4(a). A similar pattern was observed for SWS, $F(1, 25) = 9.328, p = 0.005, \eta^2 = 0.273$, also demonstrating a large effect size, see Figure 3.4(b). The average SWS was higher in the low TB condition ($M = 0.46, SD = 0.31$) compared to the high TB condition ($M = 0.311, SD = 0.187$), suggesting that a longer TB allows for more controlled and deliberate steering adjustments,

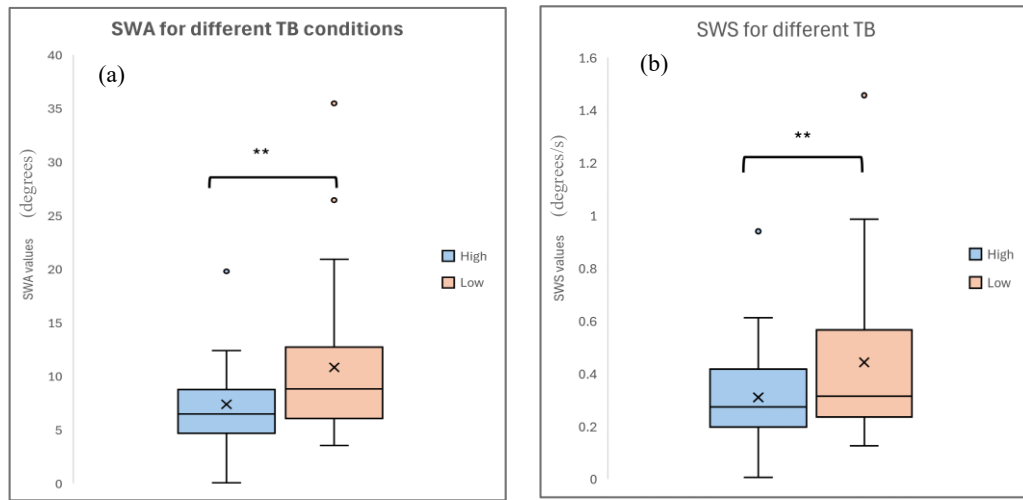


Figure 3. 4. (a) Steering Wheel angle for different TB and (b) Steering Wheel Speed for different TB.
** significant at $p < 0.005$

whereas a shorter TB may lead to more abrupt and less stable corrective actions.

In contrast, TB had no significant effect on deceleration, with $p = 0.464$ for longitudinal deceleration and $p = 0.096$ for lateral deceleration. Similarly, no significant differences were found for longitudinal SDLnP ($p = 0.408$), lateral SDLnP ($p = 0.146$), vehicle speed ($p = 0.433$),

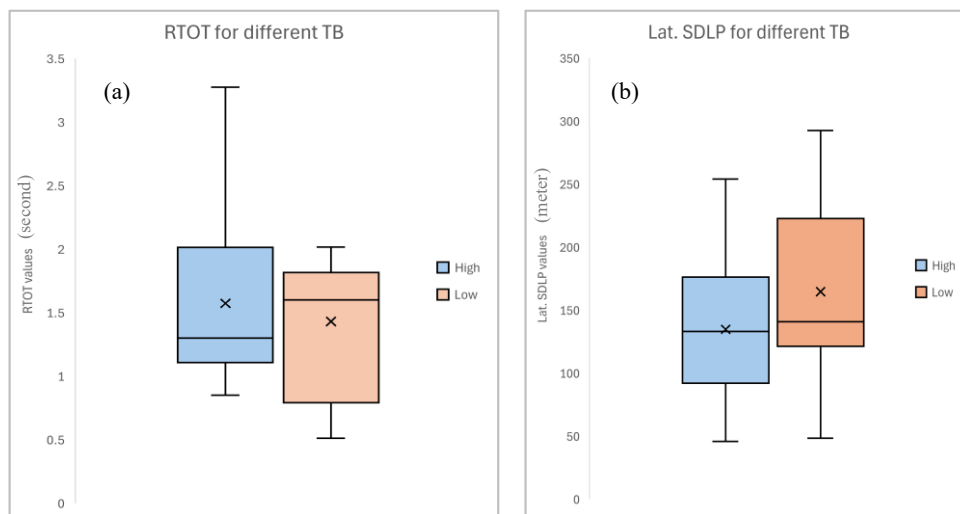


Figure 3. 5.(a) Reaction Takeover Time for different TB and (b) Lateral SDLnP for different TB.

or RTOT ($p = 0.905$), Figure 3.5(a, b). These findings indicate that while TB influences steering control, it does not significantly affect deceleration, lane stability, or reaction time after a TOR.

3.3.1.2 Effect of Age

No significant differences were observed across key driving performance metrics based on age. Specifically, age did not have a significant effect on SWA ($F(1, 25) = 0.889, p = 0.355$), SWS ($p = 0.226$), deceleration ($F(1, 25) = 1.940, p = 0.178$), longitudinal SDLnP ($p = 0.515$), or lateral SDLnP ($p = 0.176$), except for vehicle speed ($p = 0.11$). However, speed was significantly affected by age, $F(1, 25) = 9.264, p = 0.011, \eta^2 = 0.457$, with a large effect size. Some of these metrics have been shown in Figure 3.6(a, b). These results suggest that takeover performance remained relatively stable across different age groups, implying that age-related differences in cognitive or motor abilities did not significantly impact vehicle control in this context.

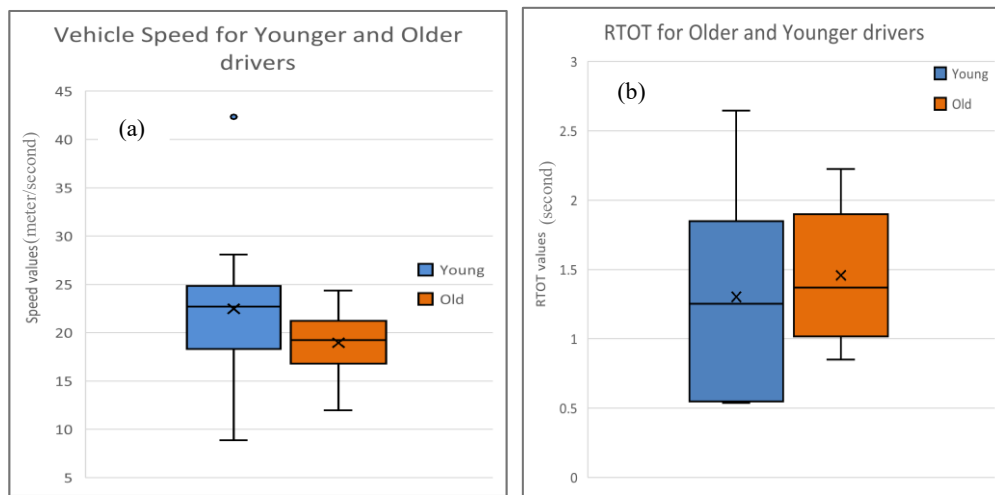


Figure 3. 6. (a) Vehicle speed for older and younger drivers and (b) Reaction Takeover Time for older and younger drivers.

3.3.2 Experiment two

The effect of TB was investigated on the above-mentioned driving behavioural metrics (see Table 3.6) and on NASA-TLX. The effect of a factor was considered significant for $p < 0.05$.

Table 3. 6. Descriptive Statistics and p -value for TOP metrics based on TB as a factor for Experiment two.

TOP Metrics	Mean High	Mean Low	p -value
Long Deceleration	1.70	2.08	0.103
Lat Deceleration	-0.15	-0.36	0.128
SDLnP	404.22	563.21	0.225
SWA	7.96	15.24	0.014*
SWS	0.37	0.56	0.041*
Vehicle Speed	20.44	24.13	0.007*
RTOT	1.32	1.19	0.674

* significant at $p < 0.005$

A significant effect of TB on SWA was observed, $F(1, 22) = 7.148$, $p = 0.014$, with a substantial effect size ($\eta^2 = 0.245$). Drivers in the low TB condition exhibited a higher SWA ($M = 15.24$, $SD = 12.45$) compared to those in the high TB condition ($M = 7.96$, $SD = 3.14$). The effect of TB was also significant for SWS, $F(1, 22) = 4.696$, $p = 0.041$, with a moderate to large effect size ($p = 0.176$). Participants in the low TB condition exhibited higher SWA and faster steering speeds ($M = 0.56$, $SD = 0.41$) compared to those in the high TB condition ($M = 0.36$, $SD = 0.17$), see Figure 3.7(a-d). The effect of TB was significant on vehicle speed as well $F(1, 22) = 10.383$, $p = 0.007$, $\eta^2 = 0.464$, where participants drove with a higher speed initially for lower TB ($M = 24.13$, $SD = 5.70$) than in the high TB condition ($M = 20.44$, $SD = 4.99$).

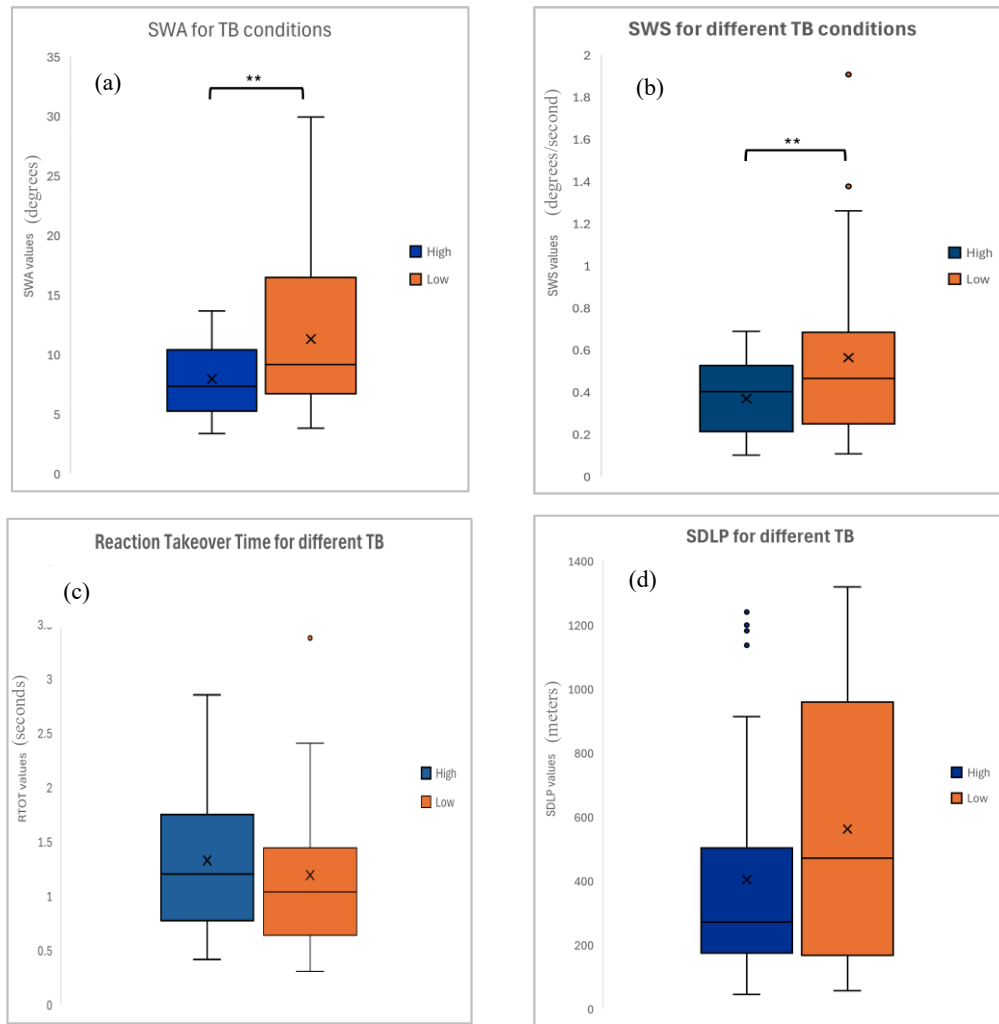


Figure 3. 7. Driving behaviour statistics after TOR (a) SWA, (b) SWS, (c) RTOT, and (d) SDLP. Low refers to 2.5 seconds of TB and High refers to 5 seconds of TB. ** significant at $p < 0.005$

No significant effect of TB was found on RTOT $F(1, 22) = 0.182, p = 0.674$. The average RTOT was 1.19 seconds ($SD = 0.75$) in the low TB condition and 1.32 seconds ($SD = 0.64$) in the high TB condition, suggesting that drivers responded consistently regardless of the available lead time. Similarly, no significant effect of TB was found on SDLP, indicating that lane-keeping stability remained unaffected by the time budget. A similar non-significant effect was found on Long. Deceleration ($p = 0.103$) and on lat. Deceleration ($p = 0.128$). Drivers across both TB conditions maintained relatively stable lateral control, suggesting that TB did not notably influence their ability to maintain lane position following a takeover request. Finally, regardless of the available takeover time, participants demonstrated similar braking responses, i.e. no significant effect of TB on resultant deceleration $F(1, 22) = 4.182, p = 0.053$.

To observe the effects of TB on workload, we conducted ANOVA analysis for NASA-TLX scores using SPSS. There was no significant effect of TB observed on the perceived workload rating, $F(1, 22) = 0.688, p = 0.602$, figure 3.8.

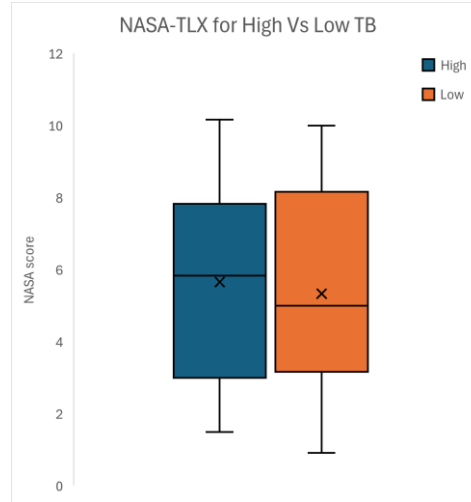


Figure 3. 8. NASA TLX for High Vs Low TB.

3.4 General Discussion

The primary objective of this study was to examine the effects of Age and TB on TOP in Level 3 automated vehicles. Specifically, we investigated how age influences TOP and whether different TB conditions affect TOP across different age groups. These findings provide valuable insights into how drivers adapt to varying levels of time pressure during takeover scenarios.

3.4.1 Effect of Age on Takeover Performance

A key research question was whether age influences TOP and whether TB affects different age groups differently. Surprisingly, the findings indicated no significant effect of age on any takeover performance metrics, including SWA, SWS, RTOT, SDLnP, and deceleration. This contradicted many studies that found poorer takeover performance in older drivers, often citing longer reaction times and slower responses to takeover notifications (Li et al., 2018, 2021; Wu et al., 2019; Peng et al., 2021). Li et al. (2021) even found that takeover time was further prolonged in older drivers of more advanced age. The only significant effect of age on TOP was on vehicle speed, where older drivers drove slower than their younger counterparts. This could be done as a strategy to compensate for their cognitive declines (Clark & Feng, 2017).

However, our findings aligned with Körber et al. (2016), Miller et al. (2016), and Favarò et al., 2019, which reported no significant differences in takeover performance between younger and older drivers. Additionally, Favarò et al., 2019 found that older drivers outperformed younger drivers in some cases, suggesting that age-related cognitive decline does not necessarily impair takeover ability. The lack of significant RTOT differences between age groups may support the

interpretation that, when provided with sufficient TB, older drivers can respond as quickly as younger ones. However, this finding should be interpreted with caution, as it may also reflect a lack of statistical power or individual variability within groups. Similarly, the absence of age effects on SDLnP and deceleration suggests that both age groups may be capable of maintaining comparable vehicle control, although other explanations – such as compensatory strategies or task familiarity – cannot be ruled out.

Regarding steering control, some studies suggested that older drivers exhibit greater variability in SWA and struggle with lane-keeping (Li et al., 2018, 2019a, 2021; Peng & Iwaki, 2020; Wu et al., 2019). However, this study accorded with other studies including Clark & Feng (2016) and Favaro et al. (2019) and found no significant differences in lane deviation or SWA between age groups. One possible explanation for the absence of age-related effects in this study was that our older participants may have had relatively high cognitive functioning and employed compensatory strategies, i.e. driving more cautiously and slowly than younger drivers as indicated by a significant difference in vehicle speed (see Figure 6(a)). Additionally, studies suggested that when given sufficient TB, older drivers could match or even exceed the performance of younger drivers due to their more deliberate and cautious approach (Khan et al., 2021).

However, these findings should be interpreted with caution, as cognitive decline varies widely among older adults (Deary et al., 2010; Hultsch et al., 2002). Furthermore, a selection bias may exist, as the study may have attracted only cognitively fit older participants, limiting the generalizability of these results to the broader ageing population. Future studies should follow different strands for ageing such as cognitive declines, individual differences, driving habits and experience, etc. to better understand how these factors influence takeover performance in older drivers.

3.4.2 Effect of Time Budget on Takeover Performance

The results from both experiments consistently demonstrated that TB significantly impacted steering behaviour, in terms of SWA and SWS, while having no significant effect on RTOT, SDLnP, or deceleration.

Drivers in the low TB (2.5s) condition exhibited greater variability in SWA and faster SWS, suggesting that time pressure caused the participants to perceive a higher sense of urgency which led to abrupt and potentially unstable corrective actions. In contrast, a longer TB (5s) allowed for more deliberate and controlled steering adjustments, possibly indicating a greater confidence in handling the situation i.e. lesser level of urgency. These findings align with previous research showing that shorter takeover times result in quicker but less precise steering inputs (Wu et al., 2020; Gold et al., 2016; Dogan et al. 2021), as drivers compensated for limited time by making more reactive steering manoeuvres. Similarly, Doubek et al. (2020) also found that participants

exhibited more abrupt steering for shorter TBs, further reinforcing the idea that TB can influence steering performance and shorter TB reduce precision in vehicle control. This abruptness in steering control was also translated to vehicle speed, where low TB noted a higher speed than high TB with a large effect size ($\eta^2 = 0.464$). However, between the two experiments, the effect of TB on vehicle speed was not consistent, likely because of differences in participants' demographics. In experiment two, the participant demographics belonged to a narrower age group of younger adults ($M = 24.79$, $SD = 7.93$), which also explains the large effect size. On the other hand, participants in Experiment One belonged to a more diverse age group ($M = 45.76$, $SD = 14.76$).

Despite the clear impact of TB on steering behaviour, RTOT remained unaffected, with drivers responding to different TB conditions within a similar timeframe. These findings were consistent with those of Du et al. (2020) who also found that TB did not influence RTOT. The non-significant effect of TB on RTOT may stem from the automatic nature of response initiation in takeover scenarios. Since RTOT reflected the initial reaction rather than full vehicle control, it likely was affected more by perceptual processing and situational awareness than by the available takeover time (Woods et al., 2015). The urgency of a TOR likely triggered a reflexive reaction, leading to similar RTOT across conditions (Shahini et al., 2023; Du et al., 2020). On the other hand, these results contradicted other findings mostly in the cases where TBs were longer (Dogan et al., 2021; Yousfi et al., 2018), possibly indicating that participants may react differently when they perceive TOR as non-urgent (Sahaï et al., 2021).

Similarly, SDLnP was not significantly affected by TB, suggesting that while TB influenced steering corrections, it did not compromise lane-keeping performance. This finding aligned with (Marfoggia et al., 2020), who also found that lateral control ability remained stable across different TB conditions, with no significant differences in SDLnP for TBs above 0 seconds. Similarly, Xu et al. (2022) found that takeover performance remained stable across varying TB conditions, even when road curvature and automation levels differed. This reinforced the idea that lane stability was relatively robust to TB variations under controlled or moderate driving conditions. However, some studies presented contradictory findings. For instance, Doubek et al. (2020) observed that shorter TBs led to greater lane deviations, particularly in high-speed or complex driving environments. This suggested that the effects of TB on SDLnP may become more pronounced in demanding scenarios. In contrast, the non-significant SDLnP resulting in this study could be attributed to the fact that both TB conditions (2.5 s and 5 s) provided sufficient time for drivers to make necessary adjustments. Even in the 2.5 s condition, where participants likely felt greater time pressure and made more abrupt steering adjustments (as evidenced by the significant effect on steering wheel angle), these aggressive inputs may have been quickly corrected or stabilized by the driver.

Additionally, deceleration (lateral and longitudinal) was not significantly impacted by TB, indicating that drivers maintained consistent braking behaviour regardless of the available takeover time. This finding aligned with studies such as Wu et al. (2022) and Doubek et al. (2020), who found that deceleration patterns in medium-traffic scenarios closely matched those of the vehicle's trajectory planner, even across different TBs. Similarly, Shahini et al. (2022) found that driving conditions had no significant effect on maximum lateral acceleration, further supporting the idea that braking and speed modulation may be more influenced by factors such as NDRTs, traffic density, road complexity, and driver workload. A possible explanation for the non-significant deceleration results in this study was that participants may have prioritized lateral control (steering) over longitudinal control (braking) during the takeover. This prioritization of lateral control over longitudinal was consistent with the situation where the task was perceived by the drivers to be easier, i.e. the time budget was sufficient for the drivers to navigate (Louw et al., 2015; Zeeb et al., 2017). As a result, drivers may have focused more on maintaining lane position and stabilizing the vehicle laterally, rather than modulating speed aggressively.

The results suggested that time budget primarily influences immediate control inputs (e.g., steering wheel angle) rather than overall stability or performance metrics (e.g., lane position, deceleration). This aligned with the idea that drivers adapt their behaviour to the available time, making more aggressive control inputs under time pressure while still achieving comparable overall outcomes. The lack of significant effects on reaction time, lane position, and deceleration further suggests that drivers can effectively compensate for shorter time budgets, or that the vehicle's automated systems provide sufficient support to maintain stability.

Given these findings, a time budget of at least 5 seconds appeared to be optimal for ensuring smoother and more stable steering inputs, as a shorter time budget (e.g., 2.5 seconds) may suffice for initiating a takeover but can lead to more erratic lateral (steering) control behaviour. These results underscore the importance of designing adaptive time budget systems for takeover requests, which can balance urgency with control stability based on real-time driving conditions.

3.5 Conclusion

This study investigated the effects of TB and age on TOP in Level 3 autonomous vehicles. The findings revealed that TB significantly influenced SWA, with Low TBs (2.5 s) leading to more abrupt and erratic control inputs compared to High TBs (5 s). However, TB did not significantly affect other performance metrics, such as RTOT, SDLnP, or deceleration, suggesting that drivers were able to compensate for low TBs and maintain overall stability. The results suggested that time budget primarily influenced immediate control inputs (e.g., steering wheel angle) rather than overall stability or performance metrics (e.g., lane position, deceleration). This aligned with the idea that drivers adapt their behaviour to the available time, making more aggressive control

inputs under time pressure while still achieving comparable overall outcomes. These results highlighted the importance of providing drivers with sufficient time to stabilize control during takeovers, particularly in scenarios requiring precise steering adjustments.

Surprisingly, age did not significantly influence any takeover performance metrics, contradicting some studies that report poorer performance in older drivers. This finding aligned with research suggesting that older drivers can match or even exceed the performance of younger drivers when given sufficient TB, potentially due to compensatory strategies such as cautious driving. However, these results should be interpreted with caution, as the study may have attracted only cognitively fit older participants, limiting the generalizability to the broader ageing population. Future research should explore how factors like cognitive decline, driving experience, and individual differences influence TOP in older drivers.

3.6 Limitations

Several limitations should be acknowledged for Experiment One. First, selection bias may be present, as the older participants likely represented a cognitively and physically healthier subset of the population, limiting the generalizability of age-related findings. Second, the study focused exclusively on critical takeover scenarios, leaving a gap in understanding how older drivers respond in non-critical or routine TORs, which may involve different cognitive demands and strategies. Designing TOR systems based solely on critical events risks overlooking important usability challenges, such as delayed situational awareness or abrupt workload transitions in more subtle handover conditions. Additionally, the outcomes may vary under different HMI designs, road geometries, or environmental conditions (e.g., urban vs. highway driving, low vs. high visibility). Finally, as with all simulator-based studies, the findings may not fully translate to real-world driving behaviour.

Across both experiments, several broader limitations apply. Real-world takeover performance is shaped by multiple interacting factors – including driver state, TOR modality, and traffic complexity – which were not manipulated in this study. This limits the ecological validity of the TB findings when applied to diverse driving contexts. Furthermore, the study lacked physiological or cognitive readiness measures (e.g., EEG, eye-tracking, or workload assessments), restricting insight into whether participants were truly prepared to resume control, particularly when behavioural metrics appeared stable. Finally, while this chapter examined the effects of Time Budget, it did not address the influence of Non-Driving-Related Tasks (NDRTs) – a factor known to modulate cognitive availability. This gap is addressed in the following chapter, which explores how engagement in different NDRTs affects takeover performance.

Chapter 4 Effect of NDRT on Takeover Performance

4.1 Introduction

In the previous chapter, it was observed that older drivers exhibited TOP comparable to that of younger drivers, despite differences in TB. The primary distinction lay in vehicle speed, with older drivers tending to drive more slowly – a characteristic often associated with ageing rather than AV technology itself. However, TOP in Level 3 AVs is influenced by multiple factors beyond age and speed such as NDRT, Gender, etc.

When a driver receives a TOR, they must swiftly disengage from their NDRT and reassume control of the vehicle. While advancements in AV technology grant drivers greater freedom to engage in NDRTs, concerns about their ability to re-engage effectively with driving remain paramount (Kyriakidis et al., 2019). Given that NDRT engagement may influence TOP, this chapter examines its impact in depth.

4.1.1 Effect of NDRT

The "out-of-the-loop" (OOTL) phenomenon refers to the state in which a human operator becomes disengaged or loses situational awareness due to over-reliance on automation (Merat et al., 2019). In this state, the operator's ability to monitor, understand, and react to changes in the system is significantly diminished.

According to Endsley and Kaber (1999), when an operator is removed from a control loop because system functions are delegated to an automated or computerized controller, the interaction between humans and the system is constrained. This limitation can lead to a decrease in the operator's awareness of system states, increasing the likelihood of errors when re-engaging with or re-assuming control over the system. This perspective has also been applied to AVs, where engagement in NDRT is found to compromise the driver's ability to monitor driving (Radlmayr et al., 2014; Wandtner et al., 2018). As a result, drivers may become OOTL, which in turn could negatively affect their driving or takeover performance (Merat et al., 2019). These decrements could be further exacerbated with the engagement in NDRT when they share the overlapping resource demands with driving according to Multiple Resource Theory, which has been comprehensively reviewed by Sebok and Wickens (2017). Numerous studies have also reported that engagement in NDRT negatively impacts driving performance following a TOR (Eriksson & Stanton, 2017; Neubauer et al., 2012; Zeeb et al., 2016). This finding is supported by several reviews and meta-analyses conducted on the subject (Zhang et al., 2019). A few studies on general transitions from system to driver showed that the higher the level of automation being utilised, the longer the drivers' reaction times to effect a successful TOR (Gold et al., 2013).

A contrary argument, derived from Stanton and Young (1997), highlights the benefits of maintaining stimulation as onboard automation increases (Stanton et al., 1997). The authors suggested that the attentional resources allocated to a given task can be redirected to new tasks thereby maintaining attentional vigilance. Later, this concept was expanded to highly automated driving (Young & Stanton, 2002). Based on this viewpoint, NDRTs can be seen as a tool to keep drivers in an “active” cognitive state that allows them to effectively take over vehicle control when necessary (Large et al., 2017). Automating the driving task reduces workload, potentially leading to an underload situation where the lack of engagement in NDRTs could impair performance and the ability to reclaim control due to a reduction in attentional resources (Stanton et al., 1997; Young & Stanton, 2002). Also, most drivers show a willingness to engage in NDRT during automated driving (de Winter et al., 2014).

In recent years, attempts have been made to reconcile these perspectives. It has been proposed that the connection between performance and cognitive workload during automated driving conforms to an inverted U-shaped function (Young, 2021). This implies that both low-load situations and high-load situations – which may render drivers OOTL – can be detrimental to driving or takeover performance as supported by the findings from the study by Ma et al. (2020). The existing research presents a nuanced picture with evidence supporting both the negative and positive impacts of NDRT engagement, as well as attempts to harmonize these views. However, gaps remain in understanding the specific conditions under which NDRTs might enhance or impair takeover performance. Therefore, this study aims to investigate the effect of NDRT (and its type) on takeover performance in L3 AVs.

4.1.2 Takeover Performance

TOP refers to the drivers’ ability to safely and effectively resume control of the vehicle after a takeover request has been initiated. This is critical in ensuring the safety of the driver and other road users when transitioning from automated to manual driving. The need for evaluating takeover performance arises from the potential risks associated with the driver being OOTL during periods of automation, which can affect their readiness and capability to manage sudden driving responsibilities.

Takeover performance is usually studied in terms of its quality and timeliness (Du et al., 2020). Takeover quality refers to the precision and control exhibited by the driver during the takeover process, ensuring that the vehicle maintains a smooth trajectory and remains within its lane. Takeover Quality is usually measured through metrics such as SWA, SDLnP, etc. Timeliness, on the other hand, pertains to how quickly the driver responds to the takeover request, which is crucial in emergencies where rapid reaction is necessary to avoid collisions or other hazards (Zeeb et al., 2016). It is usually quantified through RTOT, Deceleration, etc.

The following research question and hypothesis statement guide this investigation:

1. *Does engagement in non-driving-related tasks (NDRTs) have any effect on takeover performance?*

H1: “Active” engagement in NDRTs (Game & Video) will have a significant effect on takeover performance than “passive” engagement in the NDRT (Observation). This hypothesis is based on the premise that NDRTs could either impair or enhance the driver's ability to take over, depending on factors such as the nature of the task and the driver's cognitive workload. Additionally, the level of engagement is likely to be influenced by the number of cognitive functions involved in the NDRT. For example, the Game task requires mental, visual, and motor engagement, whereas the Video task is a ‘passive’ task that primarily requires mental and visual functions.

It is expected that participants in the Game condition, which involves simultaneous visual, cognitive, and motor engagement, will exhibit the most pronounced degradation in takeover performance. Specifically, they are likely to show greater lane deviation and less stable steering control, as the transition from the highly engaging task to manual driving may delay reorientation and motor readiness. In the Video condition, where participants are visually and cognitively engaged but not physically interacting with a task, moderate levels of impairment are anticipated. Lane deviation and steering inconsistencies may be present, though less severe than in the Game condition. By contrast, the Observation condition, which requires minimal cognitive effort and no manual involvement, is expected to produce the most stable takeover performance, reflected in minimal lane deviation and smoother, more controlled steering behaviour.

4.2 Method

4.2.1 Design

This study investigated factors associated with takeover performance using a repeated measure experimental design. The study examines two independent variables: NDRT task and Task Iteration. The NDRT task, as a repeated measure variable, had three levels: Game, Observation, and Video. Task Iteration, another repeated measure variable, indicates how often each NDRT task is repeated, with two levels: Iteration 1 and Iteration 2.

These NDRT tasks were selected as they represent naturalistic activities that participants default to when given no specific task (Detjen et al., 2020; Hecht et al., 2020). Additionally, these tasks were chosen because of their varying degree of involvement in cognitive functionalities. For example, the Game task was chosen for its ability to engage cognitive, visual, and motor functions simultaneously as has been featured in other studies (Soares et al., 2021). Subway Princess Runner (IvyMobile, 2023), an endless runner game was used where participants swipe on the screen to avoid obstacles. Similar games have been used previously and are theoretically highly engaging

due to their fast-moving pace and visual-motor demands (Rauffet et al., 2020). The video and observation tasks were selected to primarily engage cognitive and visual processing with minimal motor involvement.

This study used two video clips that depicted workplace distraction in a chemical plant and a supervisor-employee meeting as the Video Task. These tasks require a driver's auditory and visual resources (Dogan et al., 2019; Soares et al., 2021). These clips were selected for their uniqueness and their ability to maintain a consistent level of arousal throughout the duration of the video (Portron et al., 2024). Each trial in the study involved the presentation of a single video clip. The duration of clip one was two minutes and 30 seconds, and clip two lasted two minutes and 55 seconds, ensuring that neither clip concluded before the takeover request. The order of the clips was randomized to minimize any sequence effects. To ensure participants' engagement with the task, they were quizzed on the content of the videos (e.g. 'What colour suit is Sandra wearing?'), receiving a prize for their answers (Dogan et al., 2019), which was intended to maintain their focus and attention throughout the experiment. The observation task involved the participant removing their hands and feet from the controls and observing the simulator screen (Feldhütter et al., 2017; Weaver & DeLucia, 2022). This task served as a baseline to compare against the video and game tasks and to identify mental underload.

The dependent variables are mentioned in Table 4.1. When a driver needs to regain control, such as when switching from automated to manual driving, precise and measured steering inputs are essential. Therefore, analysing the SWA is crucial to assess the quality of TOP, in terms of stabilization, control, and smoothness. RTOT represents the readiness of a driver to assume control of the vehicle after a TOR. We defined RTOT as the time it takes for the driver to make 2 degrees of change in SWA from the position when the TOR was initiated (Dogan et al., 2019). SDLnP can either be measured longitudinally (along the direction of travel) or laterally (across the width of the lane). The deceleration was also recorded as longitudinal deceleration and lateral deceleration. Both the longitudinal and lateral components of both SDLnP and deceleration were combined to produce the resultant SDLnP and deceleration. We have provided table 4.1 below which details the definition of all the metrics that we have used and the levels at which these were calculated. The NASA Task Load Index (NASA-TLX) was administered after each trial to assess workload. The NASA-TLX explores the mental and physical demands of the task, time pressure, effort, performance, and fatigue (Hart & Staveland, 1988). Participants rated their response from one ('very low'/'poor') to five ('very high'/'good'). Performance answers were reversed-scored (one = 'very good') (Hart & Staveland, 1988). This study used the altered version of the NASA-

TLX, analyzing the scores of the six subscales (Hart, 2006). The NASA-TLX demonstrates high reliability and validity (Hoonakker et al., 2011).

Table 4. 1. TOP metrics and their definition.

Dependent Variables	Levels	Definition
Steering Wheel Angle (SWA) - Degrees (°)	at 2°	The standard deviation of the degree of rotation applied to the steering wheel by the driver after the TOR has been initiated.
Reaction Takeover Time (RTOT) – seconds (s)	at 2°	the time interval between the system's takeover request and the driver's initiation of control actions, i.e. steering.
Standard Deviation of Lane Position (SDLnP) - Meters (m)	at 2°	Standard deviation measures the extent to which the vehicle deviates from the center of its lane during and after the takeover.
Resultant Deceleration – (m/s ²)	at 2°	Maximum Deceleration experienced laterally by the vehicle, indicating the smoothness and stability of the takeover maneuver.

4.2.2 Participants

A total of 24 participants were recruited through snowballing and Trinity College Dublin's research system. An age range of 17-24 years was introduced to control for age effects, in line with the definition of a young driver (Department, 2022, November 24; Wu et al., 2020). Participant's inclusion criteria were having a corrected-to-normal vision, holding a full driving license, and not experiencing any major health conditions or extreme motion sickness. 12 participants identified as female and seven as male; age ranged from 19 to 24 years ($M = 22.08$, $SD = 1.47$), see Table 4.2. Participants held their license for 38.99 months on average ($SD = 22.37$), driving approximately 2.4 days per week ($SD = 2.04$). Maximum estimated driving experience since obtaining a full license was 163,100 km ($M = 25,301.24$, $SD = 39,776.26$). Participants received two research credits, where eligible, or a €20 One4All voucher. Five participants were excluded from the analysis because they either drove in an incorrect lane or did not finish the Takeover scenario.

Table 4. 2. Participants' demographics.

Participants	Count	Mean Age	License held (in-months)	Annual driver mileage (in km)
Male	7	22 (± 1.73)	41.57 (± 24.12)	4014.28 (± 3742.74)
Female	12	21.67 (± 1.30)	36.42 (± 20.63)	11520.83 (± 19430.66)

4.2.3 Materials

This study used a fixed-based, right-hand-drive driving simulator consisting of a clutch, brake, accelerator, 7-point gear stick, and force-feedback Fanatec steering wheel, see Figure 4.1. The simulator operated on SCANeR Studio software (AVSimulation, 2022). Three mounted 32-inch monitors were placed in front of the simulator. Speed (km/hr) and current gear were shown on the middle screen.



Figure 4. 1. Driving Simulator at Trinity College Dublin with a mounted Samsung Tablet.

4.2.4 Procedure

Participants were welcomed and informed consent was taken. They began with the practice scenario to familiarise themselves with the simulator. The experimenter then described the nature of AVs and task format and described the TOR. This explanation ensured that the participants understood the basic terminologies and functionalities of the AVs and allowed them to become familiar with the experiment tasks and procedures. Participants were instructed to follow the bus in their lane while maintaining a reasonable distance and a speed of 90-100 km/hr. They were instructed to manoeuvre around the lane obstacle as safely as possible.

The practice driving scenario consisted of a five-minute manual drive with limited traffic, featuring the bus ahead of the ego vehicle that needed to be followed. Participants were instructed to maintain a safe distance of approximately 3.8 meters from the lead vehicle. This distance was selected to approximate the length of a small car, ensuring a realistic minimum headway between the ego and lead vehicles. If they approached closer than this distance, an alert message was displayed (*"Too Close, Stay Back!"*), prompting them to increase the gap. In this way the distance at which the automation was engaged remained the same for all participant irrespective of the speed they were driving at. The practice session was deemed successful if participants consistently maintained a safe distance throughout the drive and they safely navigated through the obstruction in the ego lane. Once the practice session was over, the experimental scenario consisted of an eight-lane road, divided by a thick line, with mild bends (see Figure 4.2(a)). The scenery included greenery, buildings, and pedestrians. There was minimal traffic. The default

position of the ego vehicle was on the left-hand side of the road in the second lane from the centre middle line.

Participants were instructed to maintain the same ‘safe’ distance as practised during the experimental scenarios. They were also directed not to exceed a speed of 90 km/h during manual driving, while the automation maintained a constant speed of 90 km/h. The system issued an auditory TOR (a sharp beep) after some time, see Figure 4.2(b, c). The critical situation was a broken-down car in the ego vehicle’s lane, requiring an overtake to avoid a collision. Participants were clearly instructed to take control upon hearing the TOR and safely manoeuvre the vehicle. This study used a five-second Time Budget, i.e. the TOR was issued five seconds prior to transferring control to the driver. This left three seconds until collision with the obstacle, sufficient

The Experimental Scena (a)

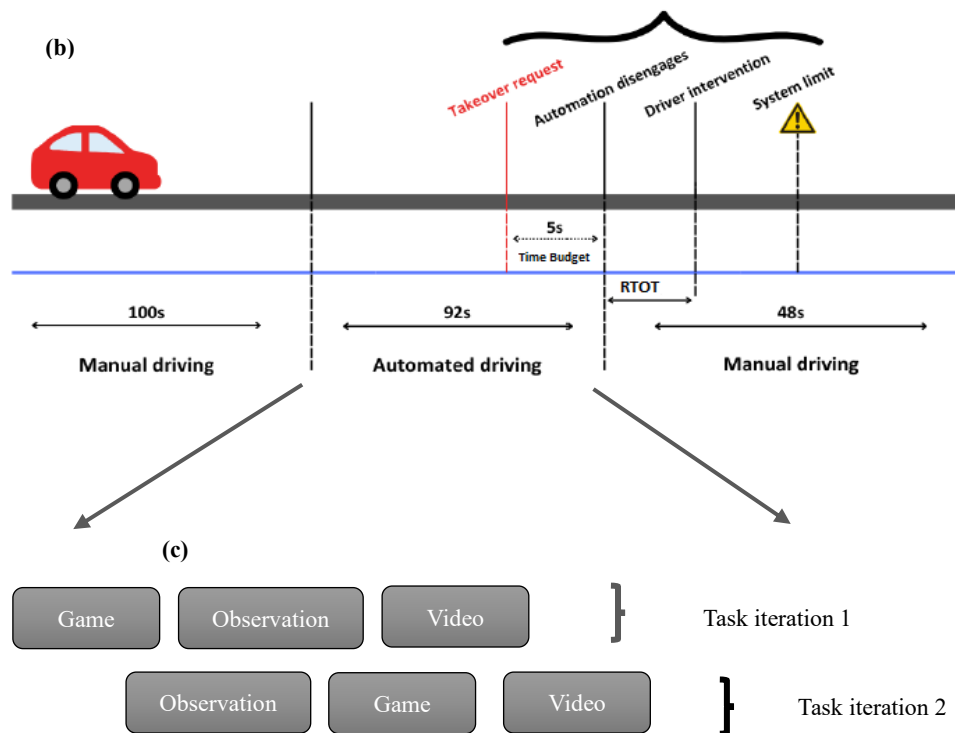


Figure 4. 2. (a) The experimental driving scenario. (b) The time points for each phase of the driving scenario. Adapted from Zhang et al. (2022). (c) NDRT tasks as Task iterations which were randomized during the experiment.

time for drivers to execute a successful takeover (Gold et al., 2013; Portron et al., 2024). Each trial lasted four minutes in total.

Each participant engaged in all three NDRTs i.e. Video, Game, and Observation tasks during automation. The Game and Video tasks were completed on the iPad. Participants completed each NDRT task twice, with task order randomly generated to negate order effects (Eriksson & Stanton, 2017). Each task was performed twice to improve data reliability and increase the likelihood of detecting an accurate performance effect (Brandenburg & Backhaus, 2016; Roche et al., 2019). Therefore, a single experimental session had six trials. Each trial began with a period of manual driving for 100 seconds before automation was engaged. Steering wheel haptic feedback and a visual alert in the top left of the screen indicated automation was on. The experimenter administered the NASA-TLX after each task trial. Participants were quizzed on clip contents after each video task. Following the six trials, participants filled out a post-experiment survey that included attitudes toward AVs (Appendix A1). The whole session lasted one hour in total. Each trial took approximately five minutes, including four minutes of driving and the time taken to administer the NASA-TLX. All participants' data were anonymised.

4.3 Results

A 3 (NDRT: Game, Video, Observation) $\times$ 2 (Task Iteration: Task Iteration 1, Task Iteration 2) repeated Analysis of Variance (ANOVA) was conducted using SPSS (v. 28.0.1.1), where NDRT and Task Iteration were within-subjects factors. The analysis aimed to assess the impact of these factors on various driving behaviour metrics, as outlined in Table 4.1. The effect of a factor was considered significant for $p < 0.05$.

Bonferroni correction was not applied in this study due to its overly conservative nature, which can increase the risk of Type II errors – particularly when outcome measures are correlated, as in this case (e.g. SWA, SWS, etc.). Instead, uncorrected p-values are reported to preserve sensitivity to potentially meaningful effects. However, we acknowledge that this approach increases the risk of Type I errors, and findings should be interpreted with caution.

4.3.1 Behavioural Measures

4.3.1.1 Effect of NDRT and Task Iteration

Similarly, the effect of NDRT was also observed on the above-mentioned metrics. The F -statistics from the ANOVA analysis are mentioned in Table 4.3.

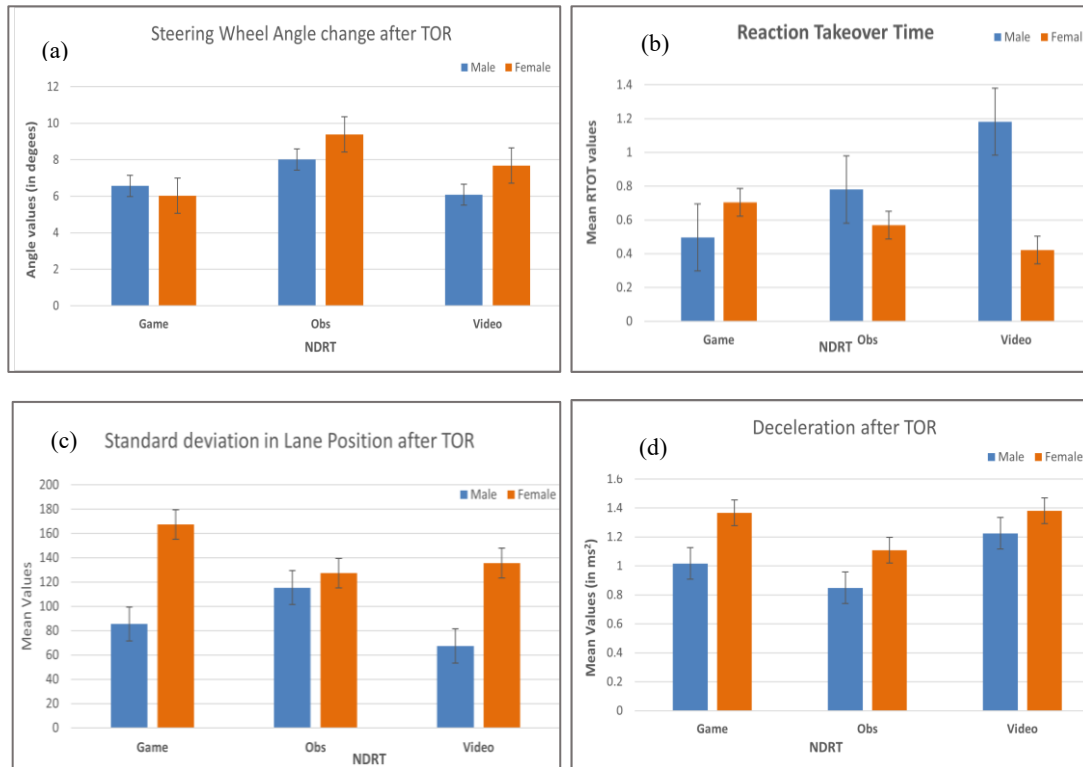


Figure 4.3. Gendered driving metrics after TOR: (a) SWA, (b) RTOT, (c) SDLnP, and (d) Deceleration.

4.3.1.1.1 Steering Wheel Angle

The factorial ANOVA analysis indicated a significant main effect of NDRT on SWA, $F(2, 34) = 3.791$, $p = 0.03$, with a $\eta_p^2 = 0.182$, which indicates a large effect size, see Figure 4.4. However, the effect of Task Iteration on SWA was non-significant, and no significant interactions between NDRT and Task Iteration were found. The mean SWA values for the different NDRTs were as follows: Game (6.30), Observation (8.70), and Video (6.89).

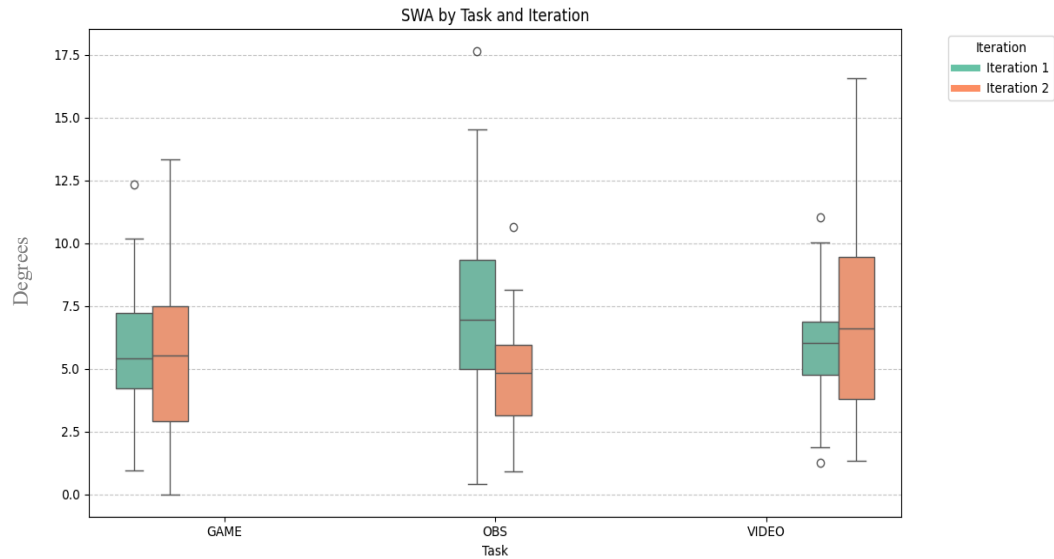


Figure 4. 4. SWA metric plot for task condition and iteration.

Table 4. 3. ANOVA Table for the effect of NDRT and Task Iterations on TOP metrics (only the significant metrics are reported).

DV/IV	NDRT (N)		Task Iteration (T)		Interaction	
	<i>F-statistics (df, df-error)</i>	<i>p-value</i>	<i>F-statistics</i>	<i>p-value</i>	<i>F-statistics</i>	<i>p-value</i>
SWA	3.791 (2, 34)	0.033** (0.182)				
RTOT			5.293 (1, 17)	0.034** (0.248)		
Resultant SDLnP					N×T: 3.892 (2, 34)	0.03** (0.207)
Resultant Deceleration			5.662 (1, 17)	0.029** (0.254)		

** significant for $p < 0.05$. N: NDRT & T: Task Iteration

4.3.1.1.2 Standard Deviation in Lane Position (SDLnP)

A significant interaction between NDRT and Task Iteration was found, with $F(2, 34) = 3.892$, $p = 0.03$, and a $\eta_p^2 = 0.207$, with a large effect size, Figure 4.5. The mean SDLnP values for the NDRTs were as follows: 126.40 for Game, 121.38 for Observation, and 101.52 for Video. Regarding Task

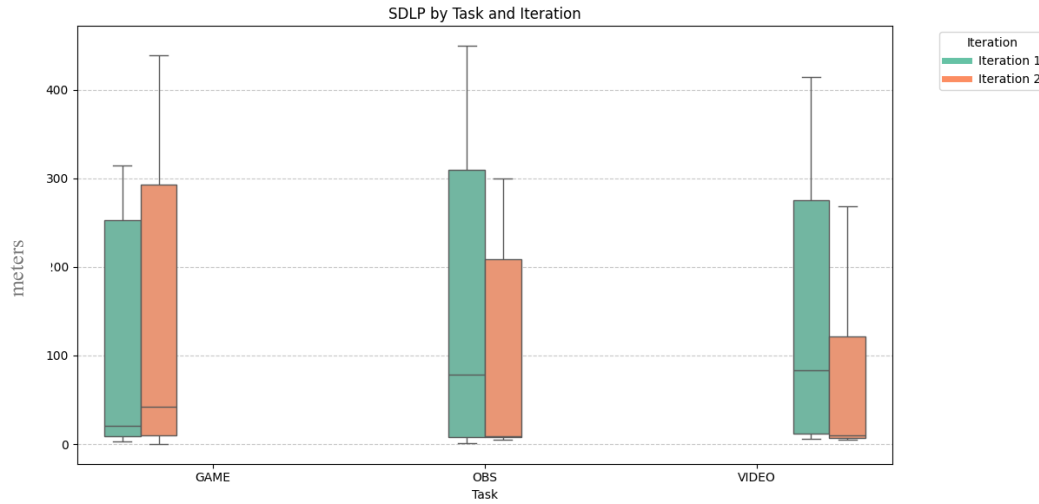


Figure 4. 5. SDLnP metric plot for task condition and iteration.

Iterations, the mean SDLnP was higher during Task Iteration 1 (133.54) compared to Task Iteration 2 (99.33).

4.3.1.1.3 Reaction Takeover Time

NDRT did not have a significant effect on RTOT. This suggests that engagement in the NDRT tasks did not meaningfully delay drivers' responses to TOR. However, task iteration had a significant effect on RTOT, $F(1, 17) = 5.293$, $p = 0.034$, with the effect size of $\eta_p^2 = 0.248$,

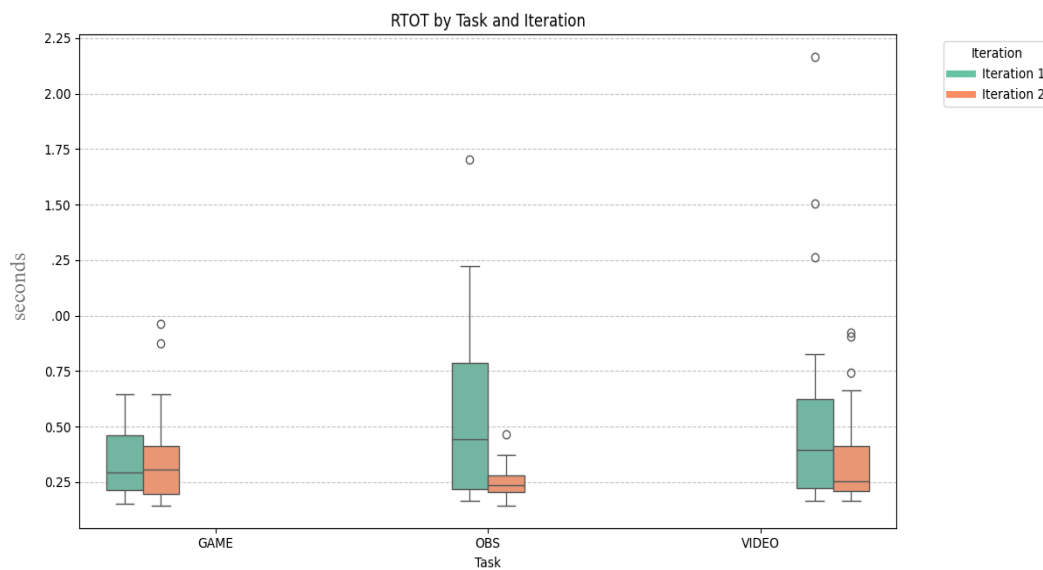


Figure 4. 6. RTOT metric plot for task condition and iteration.

suggesting a large effect size, Figure 4.6. The mean RTOT was 0.90 seconds for Task Iteration 1 and 0.43 seconds for Task Iteration 2.

4.3.1.1.4 Deceleration

The analysis revealed that NDRT tasks did not significantly affect the deceleration, suggesting that engagement in NDRT tasks did not meaningfully affect the braking performance. However, a significant effect of Task iteration, with $F(1, 17) = 5.662$, $p = 0.029$, and a $\eta_p^2 = 0.254$, indicating

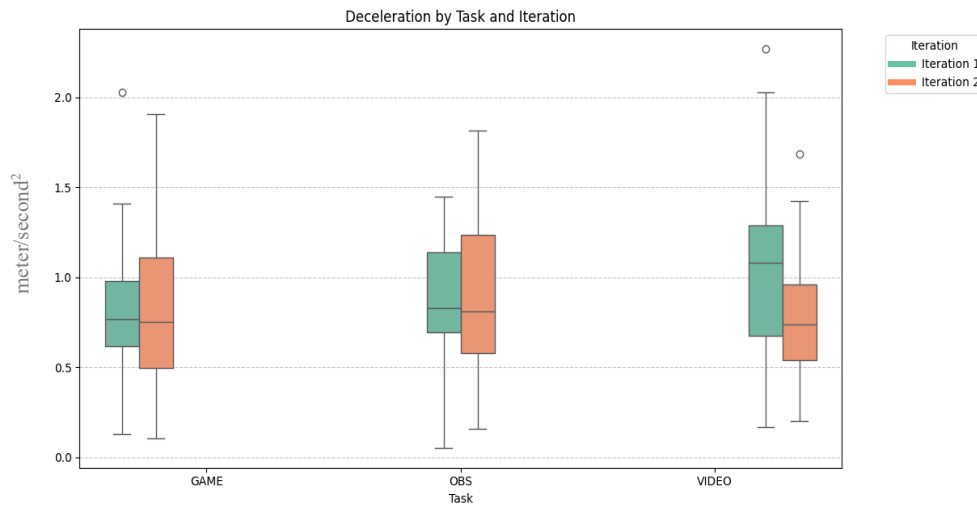


Figure 4. 7. Deceleration metric plot for task condition and iteration.

a large effect size, Figure 4.7. The mean deceleration values were 1.39 for Task Iteration 1 and 0.93 for Task Iteration 2, showing smoother braking in the second iteration.

4.3.2 Qualitative Measures

4.3.2.1 NASA-TLX (perception of workload)

Subjective Workload was estimated using 5-point NASA-TLX. Applying the factorial ANOVA, the effect of NDRT tasks on the NASA TLX did not reach significance, $p = 0.094$. Additionally, there was no effect of gender on the subjective perception of workload in AVs.

1.3.2.2 Post-Experiment Attitudes to AVs

The effect of Gender was also evaluated on a post-experiment survey where the participants were asked about their perception of reliability, trust, driving ease, safety, ease of learning and use, and usefulness, Figure 4.8(b). An ANOVA analysis was conducted and there were no significant differences ($p > 0.05$) between the perception of these items for AV context between Male and Female participants.

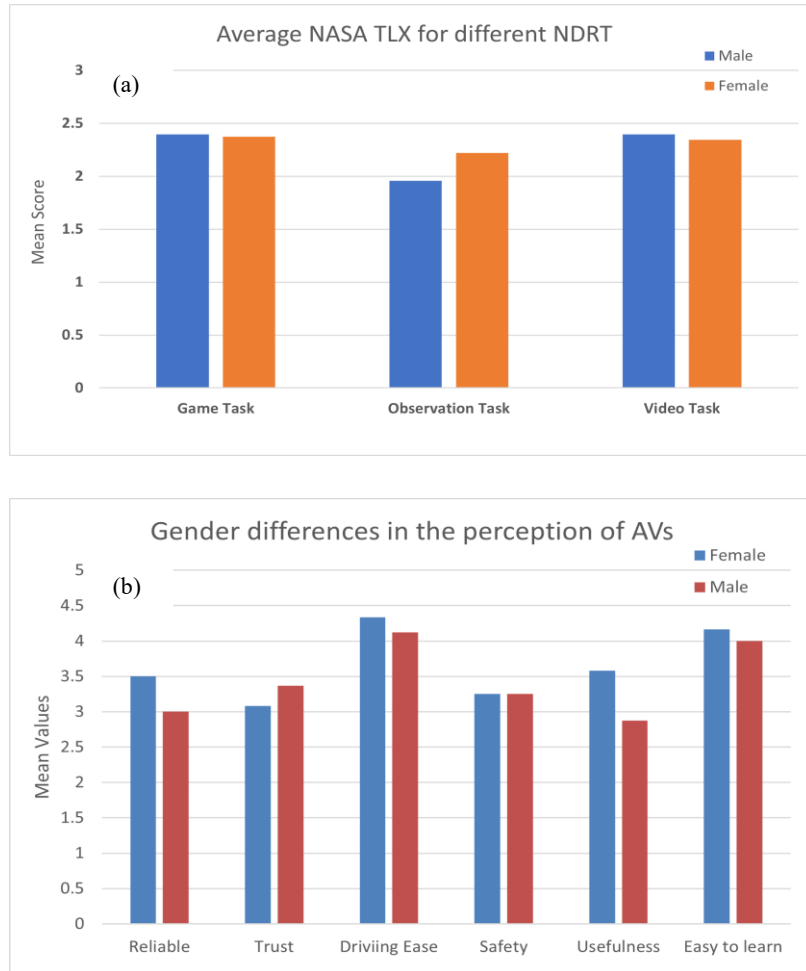


Figure 4. 8. Gender differences in: (a) subjective perception of NASA-TLX workload and (b) attitudes towards AVs.

4.4 Discussion

In this chapter, the effect of NDRT on the TOP in L3 autonomous vehicles was investigated.

4.4.1 Active Engagement in NDRTs (Game & Video) will have a significant effect on takeover performance than “passive” engagement in the NDRT (Observation).

It was hypothesized that engagement in NDRTs will have a significant effect on the TOP. This hypothesis was validated as we noted the effect of engagement in NDRT on the various metrics of takeover performance. The TOP was quantified using metrics – Reaction Takeover Time, Steering Wheel Angle, Standard Deviation in Lane Position, and Deceleration at 2 degrees of SWA.

4.4.1.1 *Steering Wheel Angle (SWA)*

Engagement in NDRT tasks significantly affected the SWA, see Figure 4.3(a) with a large effect size ($\eta_p^2 = 0.229$). These results were in line with the findings of previous studies, where they noted a significant effect of NDRT on Steering Wheel Angle (Jeong & Liu, 2019; Ou et al., 2021).

The largest SWA variability was observed for the Observation task (characterized by cognitive underload and minimal visual or motor stimulation), indicating that participants' control and stabilization were most affected during the driving environment monitoring task. Conversely, the game – which involves high cognitive, visual, and motor engagement – task showed the lowest variability overall, while the video task exhibited medium variability, see Figure 4.3(a). Improved steering performance during the Game task can be understood using the Cognitive Control Hypothesis, which posits that higher cognitive loads enhance steering performance (Engström et al., 2017). Another interpretation involves the malleable attention resource theory (Young & Stanton, 2002), suggesting that the transition from an underload state (e.g., the Observation task) to driving at TOR would be more challenging compared to transitioning from a task where cognitive-motor skills are already engaged. This would result in poorer performance in the underload condition (e.g., Observation task), while optimal engagement, as seen in more active tasks like gaming, could enhance performance (Peng et al., 2022; Shi & Bengler, 2022; Young & Stanton, 2002).

Moreover, there was no significant effect of the Task Iteration on SWA or any significant interactions between NDRT and Task Iteration, indicating that the task type, rather than the trial sequence, influenced the steering behaviour.

4.4.1.2 *Standard Deviation of Lane Position*

Driving tasks require a certain level of cognitive engagement to maintain consistent performance. When participants are underloaded, their attention may drift, leading to poorer control over the vehicle's position. This was reflected in the results from ANOVA analysis where significant interactions between NDRT and Task iterations were found in SDLnP with large effects ($\eta_p^2 = 0.207$).

In Task Iteration 1, participants showed the most variability during the Observation task in SDLnP, which decreased during the Video and Game tasks, respectively, Figure 4.5. When participants are cognitively underloaded as in Task Iteration 1, this lack of stimulation may lead to mind-wandering and disengaged gaze behaviour leading to poorer control over the vehicle's longitudinal position (Young & Stanton, 2002).

These findings align with a previous study previous research showing that under low-load conditions, drivers tend to allocate less visual attention to the road, resulting in increased lane

deviation (He et al., 2014). In the absence of a structured visual task, as in the Observation condition (which led to cognitive underload), gaze may have drifted or become unfocused, degrading steering performance. Additionally, underload can impair attentional monitoring, making drivers less responsive to subtle lane-keeping errors.

In Task Iteration 2, the Game task had the highest variability in SDLnP, while the Observation task had the second highest, and the Video task had the lowest variability, see Figure 4.4 (b). These results suggested that participants became more accustomed to the experiment paradigm and developed more trust in the automation during the second Task Iteration. This increased trust may have led to higher engagement in the Game task, with cognitive resources being heavily allocated to the NDRT, resulting in increased lane deviation when participants were asked to switch from the Game task (Peng et al., 2022; Shi & Bengler, 2022). This finding could also be linked to the competition between cognitive resources needed for the NDRT and the lane-keeping task, particularly during the takeover scenario. The lane deviation results from Task Iteration 2 are in line with the findings of the meta-analysis, where engagement in NDRT leads to more variability in lane position (Weaver & DeLucia, 2022).

While the Cognitive Control Hypothesis (Engström et al., 2017) suggests that higher cognitive load may enhance responsiveness to task-relevant stimuli, it is more applicable to non-visual tasks. For tasks involving visual-cognitive load, such as NDRTs in AV contexts, gaze behaviour plays a more direct role in modulating lane-keeping accuracy.

Moreover, a learning effect was observed between Task Iterations 1 & 2. This learning effect suggested that participants became more familiar with the experimental paradigm, potentially improving their ability to reallocate attention effectively between tasks. While this acclimatization likely contributed to improved performance during certain tasks, it may have also led to overconfidence or greater reliance on automation, influencing the variability observed during the Game task.

4.4.1.3 *Reaction Takeover Time*

An ANOVA analysis evaluated the impact of NDRT on RTOT across different scenarios, revealing that Task iteration had a significant effect on the RTOT, while, no significant effect of engagement in NDRT on the RTOT was found. For example, Lin et al. (2021) found no effect of NDRT (e.g., no task, reading the news, and watching a video) on Hands-on Time. Similar findings were reported by Dogan et al. (2019) and Vogelpohl et al. (2018) of NDRT type not affecting RTOT. Zeeb et al. (2016) also observed no significant effect of NDRT on RTOT. Together, these findings suggested that natural response tasks, such as quickly assuming steering control, may experience minimal interference from the cognitive load imposed by secondary tasks (Engström et al., 2017). Another interpretation for the non-significant effect on RTOT could be based on research that

suggests urgent TOR triggers a reflexive and automatic response which results in non-significant differences in response tasks (Shahini et al., 2023; Du et al., 2020).

The significant effect of task iteration on RTOT suggested that participants became more efficient as they progressed from Task Iteration 1 to Task Iteration 2. This finding likely reflected a learning effect, where participants became more familiar with the experiment set-up, task demands, and overall takeover paradigm, enabling them to respond more quickly in subsequent task iterations. In the context of takeover scenarios, this could imply that with practice, individuals can adapt to and better manage similar situations, leading to faster and more confident takeover responses.

4.4.1.4 *Deceleration*

A smooth deceleration indicated a controlled and stable transition, minimizing abrupt movements that could destabilize the vehicle. The analysis of deceleration revealed a significant effect of Task Iteration at a 2-degree SWA with a large effect size ($\eta_p^2 = 0.254$) suggesting that participants exhibited noticeable improvements in how smoothly they decelerated as they progressed from Task Iteration 1 to Task Iteration 2. This significant effect of task iteration indicated that, over time, participants were better able to anticipate and execute a more refined deceleration, contributing to a more stable transition. The large effect size ($\eta_p^2 = 0.254$) further underscored the presence of a meaningful learning effect, showing that the practice provided by repeated task iterations led to appreciable improvements in performance.

Deceleration seems to be more affected by the Video and Game task than the Observation task, however, the effects did not reach significance. This is in line with the results from Zeeb et al. (2017) suggested that drivers with higher cognitive tasks react more often with additional brakes compared to drivers with easier tasks. This is what we observed in the lateral deceleration where higher variability was noted for higher cognitive tasks.

4.5 Conclusion

This study demonstrated that the type of NDRT significantly impacts the takeover quality rather than the timeliness aspect of the takeover performance. These results, in general, were in line with previous studies where they found that the timeliness of the takeover performance may not be affected by the NDRT but its quality would (Dogan et al., 2019; Vogelpohl et al., 2018). Additionally, a learning effect was observed, with participants showing improved timeliness and consistency in takeover performance as they became more familiar with the task. The robustness of natural response mechanisms further enables drivers to promptly re-engage with driving, regardless of the NDRT. However, the influence of task-induced cognitive load on driving control quality highlights the need for careful integration of secondary tasks in the design of AV systems.

Integrating physiological metrics, such as EEG and eye-tracking data, with behavioural measurements could offer deeper insights into the cognitive and neurological mechanisms underpinning takeover performance. Such advancements can guide the development of more effective, personalized, and safer AV designs. Future research should also include longitudinal studies to evaluate the impact of prolonged exposure to AV systems on takeover performance. Additionally, examining demographic factors such as age, autonomous driving experience, education level, and cultural background can provide a more comprehensive understanding of driver interactions with AVs.

In conclusion, this chapter investigated the effect of engagement in NDRTs on TOP. However, this chapter did not consider the effect of these NDRT on the psychological state of the driver. In the next chapter, the effect of engagement or non-engagement in an NDRT on the driver state will be investigated, particular focus will be on induction of fatigue, measures to mitigate it and its impact on TOP.

Chapter 5 Age Differences in Fatigue Development and Perception in Level 3 Autonomous Vehicles

5.1 Introduction

In the previous chapters, we examined the TOR paradigm through key factors such as TB, age, and NDRTs. Our findings indicated that TB and NDRTs significantly influence TOP, whereas age showed minimal to no effect. However, prolonged engagement in automated driving systems can create low-workload conditions, potentially leading to reduced situational awareness, decreased motivation, diminished alertness, and ultimately, fatigue. This chapter delves into the issue of fatigue in the context of AVs, exploring its development, its impact on takeover performance, and the role of age in moderating these effects.

Fatigue, boredom and sleepiness are closely related terms. Fatigue is a psychobiological state which results from prolonged physical or cognitive exertion (Behrens et al., 2023; Boksem & Tops, 2008; Kluger et al., 2013). It encompasses both physiological and psychological components, manifesting as a decline in task engagement and overall capacity to perform. As noted by Thiffault and Bergeron (2003), fatigue involves a multidimensional experience marked by tiredness, aversion to the current task, and a decreased willingness to continue. These subjective experiences of fatigue are often compounded by the onset of boredom, particularly during monotonous or under-stimulating tasks. The attentional theory of boredom (Eastwood et al., 2012) suggests that trying to maintain attention during monotonous tasks can lead to boredom, which alters an individual's arousal state. Boredom associated with low arousal often correlates with sleepiness (Danckert et al., 2018). Monotonous tasks, such as those involving the Mackworth Clock Test, can induce sleepiness (Pickering et al., 2024). Taken together, these conceptualizations support the notion that fatigue is not solely a function of physical exhaustion, but also emerges from extended periods of cognitive underload, attentional disengagement, and motivational decline. This type of fatigue that emerges from prolonged engagement in a cognitively underload task is termed as passive fatigue (Desmond & Hancock, 2000). In contrast, active fatigue typically emerges during prolonged engagement in demanding tasks that require sustained cognitive or physical effort, leading to resource depletion over time. Both types can impair performance and situational awareness, albeit through different mechanisms, highlighting the importance of task context and cognitive load in understanding how fatigue develops and manifests.

Driver fatigue, whether resulting from boredom or other factors, has significant societal and economic consequences. Fatigued driving impairs processing speed, decision-making, and

attention, contributing to 15–25% of road crashes, particularly during nighttime or high-workload conditions (Hartley, 2004; Saxby et al., 2013). A 2009 EU analysis estimated that 10–20% of European road crashes are linked to fatigue (SafetyNet, 2009). Similarly, Langwieder et al. (1994) found that approximately 24% of crashes on German highways involved driver fatigue. These crashes often result from errors such as deteriorated steering and lane-keeping (O’Hanlon & Kelley, 1977; Philip et al., 2005). The effects of fatigue on driving performance have been compared to alcohol impairment; Williamson and Feyer (2000) reported that 17–19 hours of sleep deprivation impaired psychomotor and cognitive performance as much as a blood alcohol content of 0.5 g/L, the legal limit in many countries.

5.1.1 Fatigue in Automated Vehicles

Fatigue has long been recognized as a critical factor in impaired driving performance, but its manifestation differs depending on the level of driving automation (Endsley, 1997). In manual driving (Level 0), fatigue typically arises from prolonged cognitive and physical exertion under demanding conditions (Desmond & Hancock, 2000; Saxby et al., 2013). For instance, Åkerstedt and Kecklund (1994) showed that young drivers (aged 18–24) are especially vulnerable to fatigue-related crashes late at night, with male drivers facing double the risk compared to females. Song et al. (2017) observed that although older drivers tend to maintain stable performance, younger drivers exhibit more pronounced fatigue-related declines, particularly in speed regulation, even in conventional driving. Previous research has shown that fatigue can significantly impair cognitive and motor functions, including attention, reaction time, and hand-eye coordination (Gall, 2006; Caldwell, 2009). Moreover, fatigue has been associated with noticeable declines in driving performance – such as difficulty maintaining consistent speed and lane position, as well as slower and more error-prone decision-making – factors that collectively elevate the risk of traffic accidents (Jiang et al., 2017).

By contrast, in automated driving contexts (Level 2+), fatigue often emerges not from task overload but from cognitive underload and disengagement. The design of Level 3 AVs encourages drivers to relinquish active control, which may result in passive monitoring, mind-wandering, or boredom (Endsley, 1997; Weaver et al., 2022). This disengagement is particularly problematic in low-arousal environments like highway driving, where monotony can rapidly lead to sleepiness (Danckert et al., 2018). Empirical studies have shown that fatigue during automation can emerge quickly – Gonçalves et al. (2016) reported subjective fatigue within 15 minutes, while Körber et al. (2015) and Feldhütter et al. (2017) found physiological and eye-tracking indicators of fatigue after approximately 20–42 minutes of automated driving. Similarly, Miller et al. (2015) observed increasing fatigue when participants were not allowed to engage in non-driving-related tasks (NDRTs).

The impact of fatigue on takeover performance (TOP) in L3 AVs remains mixed. Some studies – such as Feldhütter et al. (2018), Naujoks et al. (2018), and Neubauer et al. (2012) – found no significant degradation in TOR response or control metrics, suggesting that mild fatigue may not impair short-term performance. However, other studies reported the opposite, such as, Vogelpohl et al. (2019), Hadi et al. (2020), and Pan et al. (2023) found slower reaction times, harsher braking, and poorer lane-keeping under fatigued states. W. Li et al. (2023) also reported a positive correlation between fatigue levels and RTOT. A meta-analysis by Soares et al. (2021) highlighted this inconsistency: while 75% of studies reviewed found increased RTOT under fatigue or drowsiness, only 42% of these findings were statistically significant.

Further complexity arises when considering individual differences under automated conditions. For example, Tong et al. (2024) found that novice young adult drivers (≤ 19 years) experienced more severe and rapid fatigue compared to more experienced adults, exhibiting reduced heart rate and elevated drowsiness after just 30 minutes. Wu et al. (2019) similarly found that while NDRT engagement had little impact on subjective drowsiness in older drivers, it led to reduced takeover quality. Arefnezhad et al. (2022) found that automation significantly reduced perceived workload but noted variability in how fatigue presented across age groups.

To address fatigue caused by monotony, NDRTs have been explored as potential countermeasures. Engaging in moderately stimulating activities, such as reading or interacting with touchscreens, has been shown to temporarily boost alertness. Jamson et al. (2013) and Miller et al. (2015) found that allowing drivers to engage in simple, activating NDRTs, combined with real-time feedback on the state of the automation system, could help reduce drowsiness induced by fatigue. Another study by Song et al. (2017) also highlighted the positive implications of AMT on managing fatigue, particularly in the younger adult population. Wu et al. (2019) further found that scheduled manual driving could temporarily sustain driver arousal, though its effects diminished after just 10 minutes. However, this approach is not without drawbacks. NDRTs that require significant visual or manual attention may delay responses to TORs, impairing TOP. Consequently, developing effective strategies to prevent driver Fatigue while minimizing their impact on takeover performance remains a critical challenge. Such interventions are essential for enabling drivers to maintain situational awareness and respond promptly during TORs.

Given the inconsistencies in existing findings, further research is needed to better understand how fatigue influences takeover performance in L3 AVs. Understanding these relationships is critical for the safe deployment of automated driving systems, particularly those requiring intermittent human intervention.

To address this gap, the current chapter explored the following research questions:

5.1.2 Research Questions

- i. Can Fatigue be induced by experiencing a low workload task in autonomous vehicles?
- ii. How does Age interact with Fatigue in affecting TOP?
- iii. How does Fatigue affect TOP in comparison to NDRT's effect on TOP?
- iv. Can engagement in NDRT be used to mitigate Fatigue levels?

5.2 Methodology

5.2.1 Participants

Participants were recruited as part of the FotoNation/Tobii drowsiness study. The required sample size was determined using G*Power 3.1.9.4, based on a two-tailed alpha of 0.05, a power of 0.8, and an anticipated medium effect size ($\eta_p^2 = 0.06$). This effect size was chosen as it aligns with prior research findings and represents the smallest effect size of interest (Riesthuis, 2024).

A total of 34 participants (17 female, 14 male, 1 non-binary) were divided into two age groups: younger adults (18–49 years, $n = 22$) and “50+ age older adults” (50+ years, $n = 12$), following the categorization by Arefnezhad et al. (2022), see Table 5.1. The overall age range was 18 to 69 years ($M = 40.6$, $SD = 18.4$). The inclusion of older adults aimed to explore potential age-related differences in fatigue and takeover performance. The sample was designed to ensure demographic representation for broader generalizability to future AV users (see Figure 5.1). Efforts were also made to promote gender inclusivity.

Participants were recruited by Tobii Galway for a 2.5-hour driving experiment using a recruitment agency, academic contacts, and snowballing. They were compensated with €50 for their time and participation. Exclusion criteria were a history of major heart condition conditions, severe motion sickness in a virtual environment, Tinnitus, or any other health condition that obstructed them from being fitted with wearable sensors.

Table 5. 1. Participants' demographic data.

Age Group	Female	Male	Non-binary	Total
Younger Adults (18–49)	10	11	1	22
50+ age Older Adults (50+)	5	7	0	12
Total	15	18	1	34

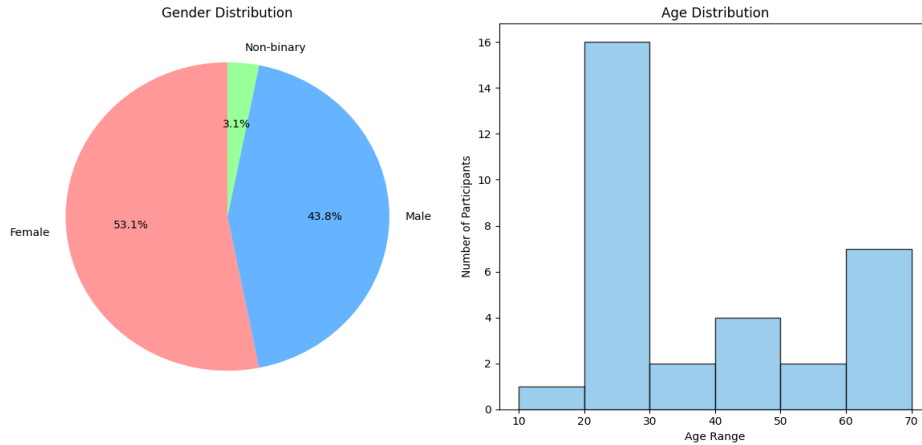


Figure 5. 1. Demographic distribution of Participants.

5.2.2 Materials

The experiment was conducted using the high-fidelity FotoNation/Tobii in-house dynamic driving simulator (DS), depicted in Figure 5.2. This state-of-the-art simulator featured a fully equipped car interior mounted on an actuated platform, designed to replicate real-world driving dynamics. It included a force-feedback-enabled steering wheel, pedals, and gearstick, which enhanced the realism of the driving experience. To provide an immersive visual environment, the simulator was equipped with three 49-inch monitors arranged to create a seamless, wide-field display that rendered high-resolution driving scenes in front of the driver. Additionally, a smaller, elevated monitor was strategically positioned to display parallel tasks without obstructing the primary field of view.



Figure 5. 2. High Fidelity dynamic driving simulator at Tobii.

The simulator's hardware and simulation software were managed by an integrated system: one computer oversaw the DS operation and simulation environment, while a separate system handled the collection of multi-modal data, including video, audio, and other sensor outputs. The

simulation environment was powered by SCANeR Studio software (AVSimulation, 2022), which enables precise control of driving scenarios and data acquisition.

5.2.3 Experiment Design

This study employed a mixed-factorial design to investigate the effects of fatigue, workload, and TOP. The between-subjects variable was age, with participants grouped into two categories: young drivers (aged 18-49) and “50+ age older drivers” (aged 50+). The within-subjects variable was the two task conditions: Drive Monitoring and NDRT, presented in a randomized order to control for potential order effects.

The dependent variables included both objective and subjective measures. Behavioural data were collected from the driving simulator, capturing key performance metrics such as steering responses, braking behaviour, and lane position. Subjective assessments were conducted using the NASA Task Load Index (NASA-TLX) to evaluate workload and the Karolinska Sleepiness Scale (KSS) to measure fatigue levels. Additionally, video analysis was incorporated to provide an observational measure of fatigue, identifying signs such as prolonged eye closures and head nodding. This multimodal approach ensured a comprehensive evaluation of fatigue and its impact on takeover performance in Level 3 automated vehicles.

5.2.4 Experimental Condition

The study included two experimental conditions: NDRT Condition and Drive monitoring Condition. Initially, in the practice trial, participants performed a manual driving task, where they were required to navigate around a broken vehicle that appeared ahead of them in their lane while trying to follow a lead vehicle, see Figure 5.3. This was to ensure they were familiarized with the driving simulator set-up, with experimental paradigm, tasks, and conditions.



Figure 5. 3. Driving virtual environment of the simulator from SCANeR software.

In the NDRT Condition, participants initially drove manually for 5 minutes, after which the automation was engaged for around 14 minutes. During this automated driving period, they were instructed to play a card-matching game on a Samsung tablet, which was affixed to the driving setup using an adjustable arm. This NDRT was chosen to provide cognitive engagement and help

counteract drowsiness while the vehicle operated autonomously (Coyne et al., 2023). Moreover, this task was intentionally selected over a more complex working memory challenge, such as the N-back task, to prevent the onset of task-related fatigue, which can occur when sustained attention is required for a highly demanding cognitive activity (Pickering et al., 2024). The game itself featured a 10×10 grid of cards, all initially placed face-down. Participants could tap on a card to flip it over and reveal its image, but only one card could be viewed at a time. The objective was to remember the positions of specific cards and systematically match all pairs until the entire set was completed. When the automation reached its operational limit, it issued a TOR via a beep, giving the driver a 3-second time budget to assume control.

In the drive monitoring condition, participants initially drove manually for 5 minutes, after which automation was engaged. While the automation was engaged, participants were instructed to simply observe the automated driving environment displayed on the three screens, without engaging in any active driving or tasks. For both the NDRT and Drive monitoring conditions, a TOR was initiated after 14 & 40 minutes respectively during automated driving. When a TOR was triggered, a beep sound would signal the participants that they had 3 seconds to resume manual control of the vehicle and safely navigate around the broken vehicle scenario.

5.2.5 Takeover Performance

Takeover performance refers to a driver's ability to safely and effectively regain control of a vehicle following a TOR and is evaluated through timeliness and quality (Zeeb et al., 2016). Takeover quality reflects the driver's precision and control during the manoeuvre, assessed through metrics such as SWA, SWS, and SDLnP. Timeliness refers to the quickness in the driver's response to a TOR, with metrics like RTOT, braking, and deceleration performance being critical for collision prevention in emergencies.

In this study, several dependent variables were used to evaluate both aspects of takeover performance. SWA and SWS measured vehicle stabilization and smooth manoeuvring, while RTOT, defined as the time to initiate a 2-degree steering change after a TOR (Dogan et al., 2019), captured response speed. Braking performance assessed the driver's ability to decelerate effectively, and SDLnP (both longitudinal and lateral) evaluated driving precision.

Additionally, Subjective workload was evaluated using the NASA-TLX scale, and fatigue/drowsiness was assessed using the KSS. A full list of TOP is presented in Table 5.2.

Table 5. 2. Takeover Performance metrics.

Dependent Variables	Levels	Definition
Steering Wheel Angle - Degrees (°)	at 2°	The standard deviation of the degree of rotation applied to the steering wheel by the driver after the TOR has been initiated.
Steering Wheel Angle - Degrees (°)	at 2°	The rate of change in the steering wheel angle (SWA) over time, indicating how quickly the driver adjusts steering after the takeover request (TOR).
Reaction Time - Seconds (s)	at 2°	the time interval between the system's takeover request and the driver's initiation of control actions, i.e. steering.
Standard Lane Deviation - Meters (m)	at 2°	Standard deviation measures the extent to which the vehicle deviates from the center of its lane during and after the takeover.
Deceleration - (m/s ²)	at 2°	Maximum Deceleration experienced laterally by the vehicle, indicating the smoothness and stability of the takeover maneuver.
Brake - Newtons (N)	at 2°	The braking force applied, measured in Newtons (N).

5.2.6 Procedure

The procedure began with an initial screening phase where participants were contacted by an experimenter who explained the study's structure, objectives, and inclusion criteria. Following this, participants received a detailed written description of the experiment along with an informed consent form. This document outlined the experiment design, types of sensors to be used, participants' rights (including the right to withdraw at any point), and GDPR-compliant data protection measures. To ensure participants started the experiment in a rested state, they were instructed to obtain at least 10 hours of healthy sleep the night before the session and to abstain from consuming caffeinated beverages on the day of the experiment. The experiment was carried out in two separate sessions in the morning session or in the afternoon session, with these preconditions verbally verified upon arrival at the facility. Upon arrival, participants were given an informed consent form to sign. They were introduced to the driving simulator setup, including its controls and hardware, and were provided with a detailed explanation of the tasks to be performed in each experimental condition. The task in each experimental condition was to follow a lead vehicle at a safe distance and safely manoeuvre around any obstructions in their lane when necessary. The experimenter explained to the participants that automation would carry out longitudinal and lateral control, and they should take their hands off the steering wheel and their feet off the pedals while automation is engaged, as shown by the “automation ON” sign on the screen. They were required to regain control as fast as they could when there were situations that the automation could not solve, during which situation a TOR would be issued with a time budget of 3 seconds.

To familiarize participants with the driving simulator, a practice session was conducted. The practice scenario consisted of a five-minute manual driving task in a low-traffic environment, during which participants were instructed to follow a lead vehicle at a speed of 90 km/h while maintaining a safe distance of 3.8 meters. If the participants reduced the distance below this threshold, an alert message prompted them to increase the gap. The practice session was deemed successful if participants consistently maintained a safe following distance and successfully navigated an obstruction in their lane.

Following the practice session, participants began the experimental phase, which took place in a simulated environment featuring an eight-lane road with mild bends, greenery, buildings, pedestrians, and minimal surrounding traffic, as illustrated in Figure 5.3. The ego vehicle was positioned in the second lane from the centre on the left-hand side of the road. Participants completed both experimental conditions (NDRT and Drive monitoring) in a randomized order to minimize the effects of presentation bias or task familiarity. For the NDRT and Drive monitoring conditions, participants responded to TORs as described earlier, ensuring consistent exposure to the automated driving environment and takeover scenarios across all conditions.

5.3 Results

5.3.1 Descriptive Statistics

Table 5.3 provides the descriptive statistics for each of the behavioural measures.

Table 5. 3. Descriptive statistics of each of the TOP metrics.

Variable	Condition	Mean	SD	Min	Max
Long. Deceleration	Fatigue	2.85	2.19	-0.32	6.41
	NDRT	2.80	2.29	-0.30	6.48
Lat. Deceleration	Fatigue	0.004	0.93	-2.67	2.03
	NDRT	0.13	0.54	-0.31	2.19
Long. SDLNP	Fatigue	39.64	14.97	4.34	60.51
	NDRT	50.20	19.03	0.41	97.87
Lat. SDLNP	Fatigue	24.33	10.45	0.93	48.51
	NDRT	28.63	16.46	0.1	79.53
SWA	Fatigue	18.46	12.08	0.38	41.37
	NDRT	10.19	8.89	0.19	30.41
SWS	Fatigue	1.05	1.03	0.01	3.91
	NDRT	0.59	0.68	0	3.09
RTOT	Fatigue	3.12	1.63	0.9	6.95
	NDRT	3.38	1.42	1.6	7.0
Brake	Fatigue	77.87	90.08	0	268.79
	NDRT	87.54	104.98	0	309.19

5.3.2 Statistical Analysis

A mixed factorial ANOVA was conducted on the TOP measures using SPSS, with Age as a between-subjects factor and Task Condition as a within-subjects factor. Descriptive statistics for these TOP measures have been provided in Table 5.3. There was no significant interaction between Age and Task Conditions. Therefore, only the main effects were reported below.

Bonferroni correction was not applied in this study due to its overly conservative nature, which can increase the risk of Type II errors – particularly when outcome measures are correlated, as in this case (e.g. SWA, SWS, etc.). Instead, uncorrected p -values are reported to preserve sensitivity to potentially meaningful effects. However, we acknowledge that this approach increases the risk of Type I errors, and findings should be interpreted with caution.

5.3.2.1 Effect of Age

The influence of age on TOP metrics listed in Table 5.2 was analysed to assess potential differences between younger and older participants. Age had no significant effect on any of the TOP measures. Given that RTOT is a commonly used TOP metric, it is highlighted here, with no significant difference observed between age groups ($p = 0.766$). Other TOP metrics examined also showed no significant age effects.

5.3.2.2 Effect of Task Conditions on Takeover Performance

The impact of task conditions (Drive monitoring vs. NDRT) on various TOP measures was assessed as the main effects. The results showed differing effects across the variables.

Longitudinal Deceleration: There was no significant effect of task condition on longitudinal deceleration, $F(1, 28) = 0.142$, $p = 0.71$, $\eta^2 = 0.005$, indicating that neither drive monitoring nor NDRT conditions significantly affected deceleration in the longitudinal direction.

Lateral Deceleration: A significant effect of task condition was observed on lateral deceleration, $F(1, 28) = 5.720$, $p = 0.024$, $\eta^2 = 0.170$, see Figure 5.4. This suggests that task condition had a significant impact with NDRT showing lower values. It is important to note that outliers are likely to be influencing these results.

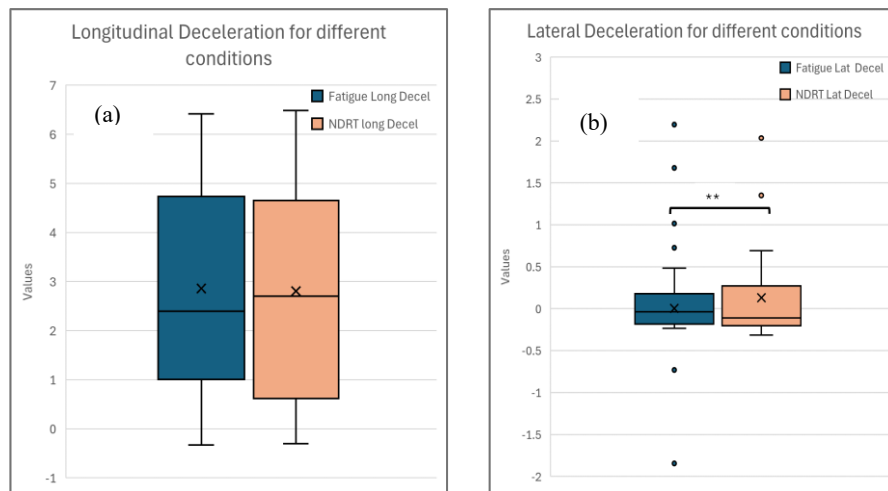


Figure 5. 4. Longitudinal and Lateral Deceleration for Fatigue and NDRT conditions. ** significance at $p < 0.05$

Longitudinal Standard Deviation in Lane Position (SDLnP): The task condition significantly affected longitudinal SDLnP, $F(1, 28) = 12.93$, $p = 0.001$, $\eta^2 = 0.332$, indicating a large effect of task conditions on SDLnP in the longitudinal direction.

Lateral Standard Deviation in Lane Position (SDLnP): A significant effect was found for task condition on lateral SDLnP, $F(1, 28) = 5.499$, $p = 0.029$, $\eta^2 = 0.208$, see Figure 5.5. This indicates that Drive monitoring and NDRT conditions influenced the lateral SDLnP, with a large effect size.

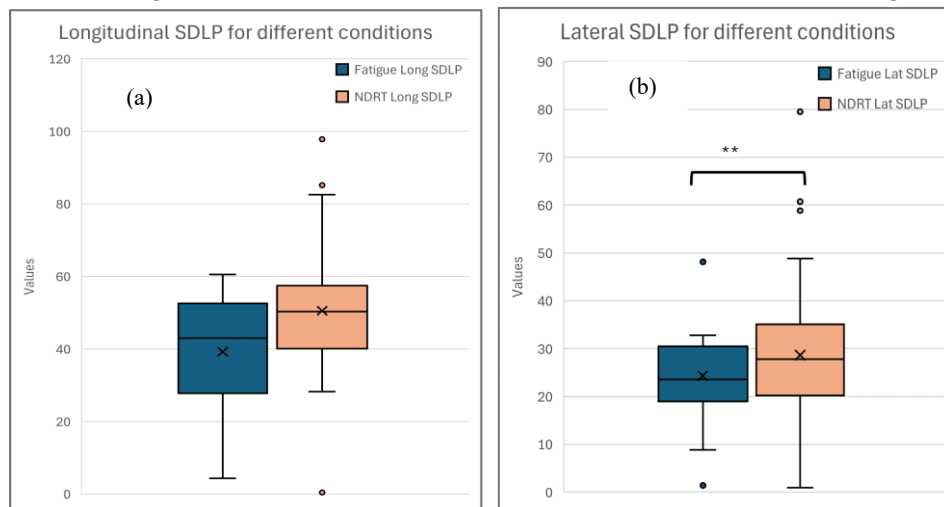


Figure 5. 5. Longitudinal and Lateral SDLnP for Fatigue and NDRT conditions. ** significance at $p < 0.05$

Steering Wheel Angle (SWA): The task condition significantly affected Steering Wheel Angle, $F(1, 28) = 5.964$, $p = 0.024$, $\eta^2 = 0.221$, demonstrating that Drive Monitoring and NDRT conditions had a large effect on steering control, see Figure 5. 6(a).

Steering Wheel Angle Speed (SWS): The effect of task condition on Steering Wheel Angle Speed was significant, $F(1, 28) = 4.796$, $p = 0.038$, $\eta^2 = 0.167$, indicating that both Drive monitoring and NDRT conditions influenced the speed at which participants turned the steering wheel, see Figure 5.6(b).

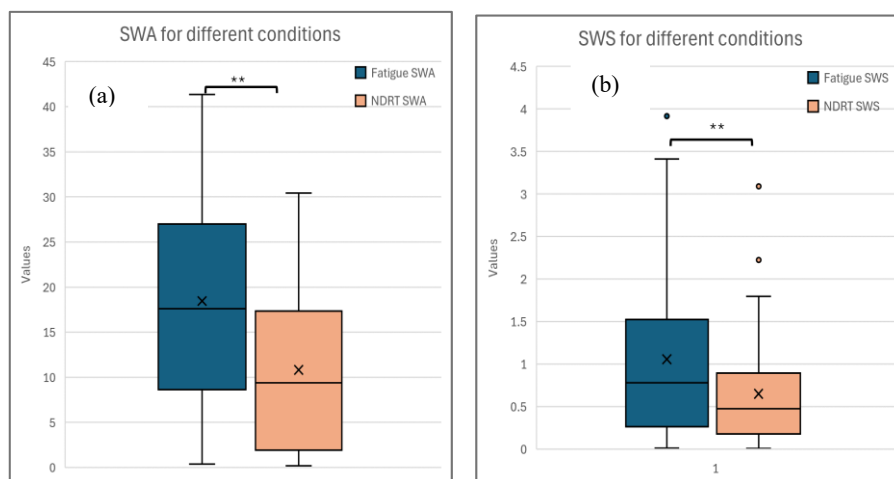


Figure 5. 6. Steering Wheel Angle and Steering Wheel Angle Speed for Drive Monitoring and NDRT condition. ** significance at $p < 0.05$

Reaction Takeover Time (RTOT): No significant effect was found for task condition on Reaction Takeover Time, $F(1, 28) = 1.218$, $p = 0.29$, $\eta^2 = 0.086$, suggesting that neither Drive monitoring

nor NDRT had a significant impact on the time taken to initiate a steering response during a takeover request, see Figure 5.7 (a).

Brake: No significant effect of Task Condition was found on braking performance, $F(1, 28) = 0.142$, $p = 0.29$, $\eta^2 = 0.005$, indicating that braking responses were similar across both Drive monitoring and NDRT conditions, figure 5.7(b). This suggested that task engagement did not influence participants' braking behaviour during the takeover event.

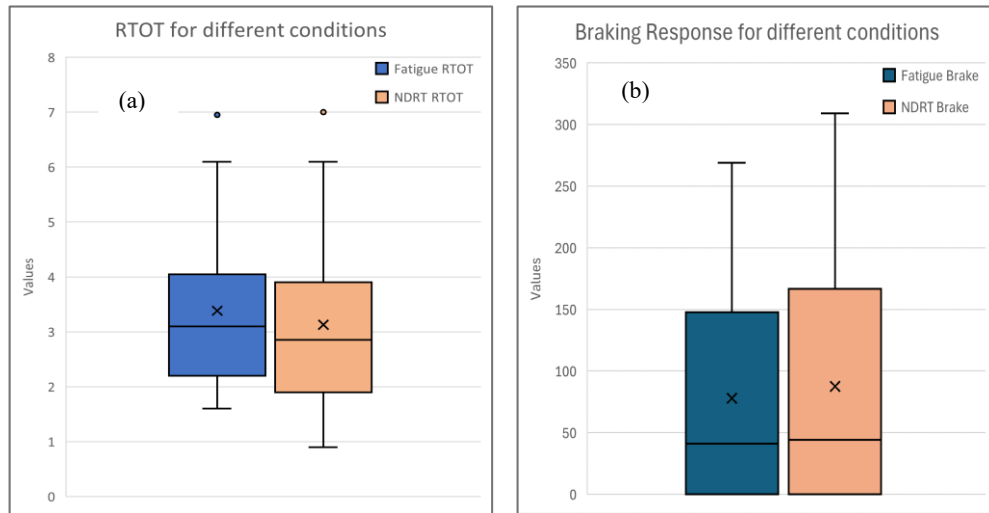


Figure 5. 7. Reaction Takeover Time and Brake for Fatigue and NDRT conditions.

5.3.2.1 Effect of Age and Task Condition on Driver Workload (NASA-TLX) and Fatigue Levels (using KSS scores)

A one-way between-groups ANOVA was performed on the reported KSS values for the NDRT and Drive monitoring tasks. No significant interaction was found between the age and task condition, so the main effects were reported.

5.3.2.1.1 Effect of Age on workload and fatigue levels

The analysis revealed no significant effect of age on KSS, $F(1, 28) = 0.239$, $p = 0.628$, indicating that age did not influence subjective sleepiness levels during the task. In other words, both younger and older drivers reported similar perceptions of sleepiness. Interestingly, age had a significant effect on the NASA-TLX scores, $F(1, 28) = 14.343$, $p < 0.001$, $\eta^2 = 0.339$, figure 5.8(a, b), with older drivers reporting higher workload across both task conditions.

5.3.2.1.2 Effect of Task Conditions on workload and fatigue levels

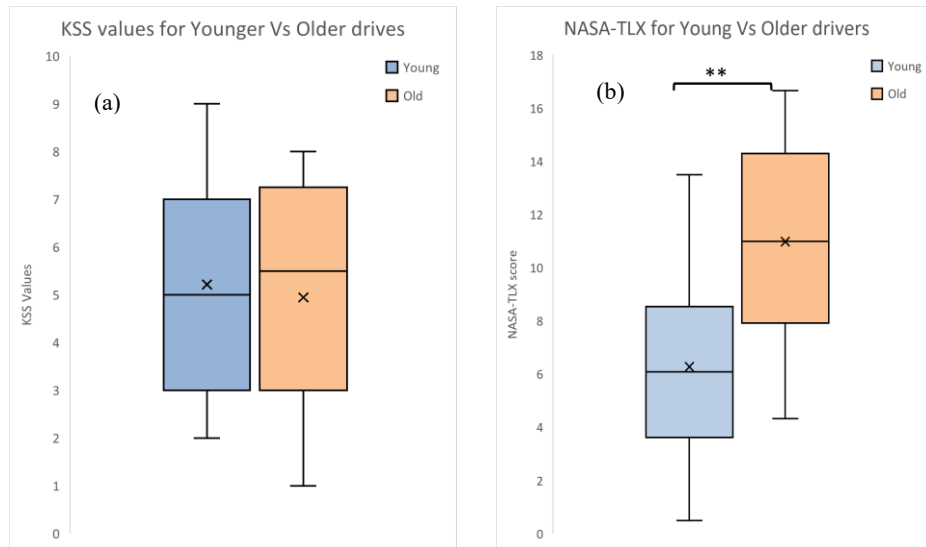


Figure 5. 8. Self-reports for Younger and Older adults in different conditions for (a) KSS and (b) NASA-TLX. ** significant for $p < 0.05$

The same model was applied to the NASA-TLX scores (perceived workload) and KSS values for both conditions. Figure 5.9 presents the frequency distribution of self-reported KSS values under both conditions. The majority of participants reported experiencing significant sleepiness, with the highest frequency observed in the "Very sleepy; significant effort to keep alert, fighting sleep" category. Thus, a significant effect of task condition was observed on KSS, $F(1, 28) = 67.155$, $p < 0.001$, $\eta^2 = 0.706$ with a large effect size, where drive monitoring condition produced most fatigue effects, see figure 5.10(a). This indicated that the drive monitoring task led to significant fatigue, with 80% of participants reporting signs of drowsiness under the drive monitoring condition. In contrast, the effect of task condition was not significant on NASA-TLX scores, $F(1,$

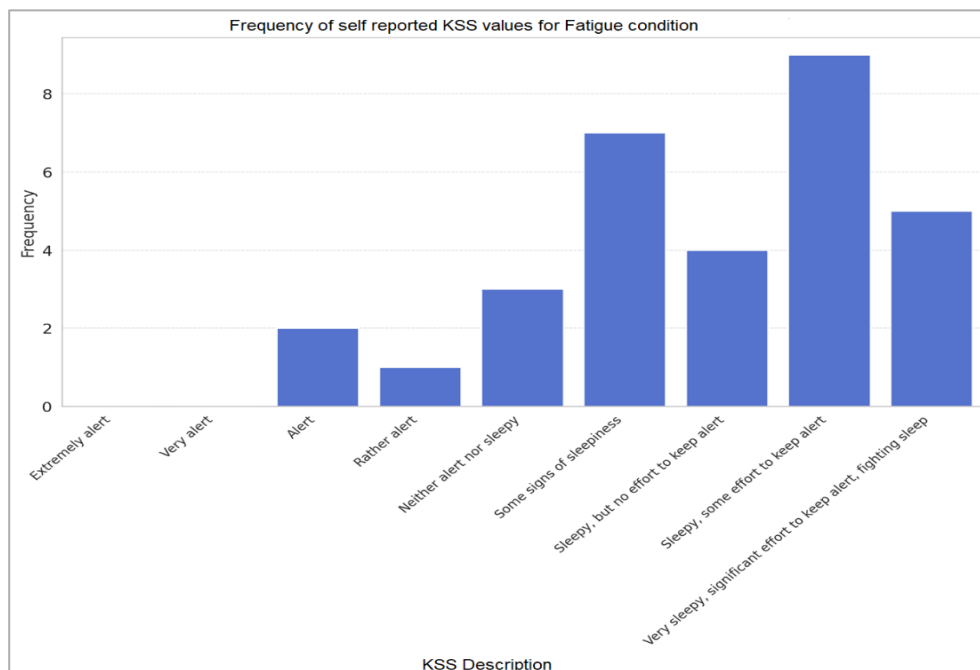


Figure 5. 9. Frequency distribution of reported KSS levels with their corresponding description.

28) = 0.429, $p = 0.518$, figure 5.10(b). This suggested that the drive monitoring task did not make participants feel more mentally fatigued or challenged compared to the NDRT task.

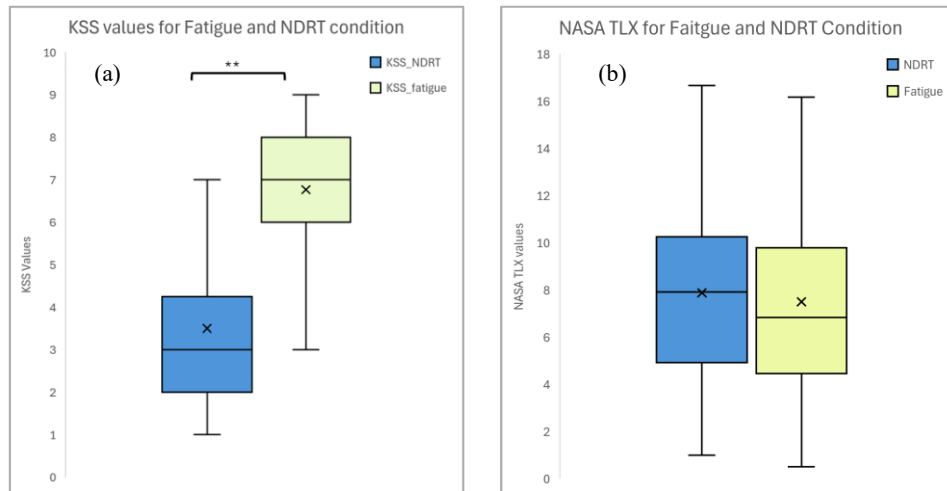


Figure 5. 10. Self-reported (a) KSS for NDRT and Fatigue conditions, and (b) NASA-TLX for fatigue and NDRT condition. ** significant for $p < 0.05$

5.4 Discussion

This study investigated the impact of prolonged automation, specifically in L3 AVs on driver drowsiness. It also investigated the relationship the drowsiness has with NDRT, driver age and its impact on TOP. The results provide important insights into these relationships, which are discussed below in the context of the research questions.

5.4.1 Effect of Age on Takeover Performance

This study found no significant difference between the two age groups investigated on objective performance measures such as RTOT, SDLnP, SWA, or SWS. These findings align with previous research, including Arefnezhad et al. (2022), who also found no significant age-related differences in driver performance. While the results suggest that younger and “50+ age older drivers” may perform similarly in takeover scenarios within automated systems, this interpretation should be treated with caution. The absence of statistically significant age effects does not conclusively demonstrate equivalence, as the findings may also be influenced by limited statistical power or other unaccounted-for moderating variables that could obscure subtle age-related differences in performance. The findings of this study align with prior research on age-related differences in driving performance, such as Song et al. (2017), which also reported no significant differences in SDLnP between age groups. This consistency reinforces the conclusion that both younger and older drivers exhibit comparable TOP. However, Song et al. (2017) did identify age-related differences in speed variability, suggesting potential discrepancies influenced by study design, experimental tasks, or contextual factors such as automation level. A key distinction between the two studies lies in the nature of the takeover scenarios. While the present study focused on TORs

requiring an urgent driver response, Song et al. (2017) examined "normal" (i.e. manual) driving conditions. Additionally, methodological differences may have contributed to the observed variations. This study employed a repeated-measures design with a randomized condition order, whereas Song et al. (2017) used a fixed sequence, transitioning from monotonous driving to baseline and attention-maintenance tasks. These differences underscore the need for further research into task-specific variables that may influence takeover performance across different age groups in automated driving contexts. For instance, factors such as TOR urgency, road context, control modality, and prior engagement (e.g., NDRTs vs. passive monitoring) can differentially tax cognitive and motor systems, especially in older adults. In low-urgency or predictable scenarios, age-related differences may become more pronounced due to the availability of time as noted in (Li S. et al., 2019). However, in high-urgency or cognitively complex transitions, older drivers may perform comparably to younger ones due to reflexive or automatic processes involved as noted in chapter 3. Thus, task-specific variables can moderate the relationship between age and TOP, helping to explain why some studies (like Song et al., 2017) find age-related impairments while others (like this) do not.

Interestingly, there were no significant age effects on KSS values – suggesting similar perceived fatigue levels between age groups. This was in contrast with what has been found by Tong et al. (2024) that younger drivers experienced fatigue quicker. It is important to note that KSS values are self-reports may not always align with objective measures, such as heart rate variability, etc. Moreover, for perceived workload older drivers reported higher perceived workload levels compared to younger drivers, see Figure 11(b). This significant difference was consistent with existing research (Soares et al., 2020), which suggested that older drivers may experience increased cognitive effort when performing demanding tasks such as TOR, even when objective performance remains similar (Cooper et al., 2020; Strayer et al., 2015). The finding that older drivers perceive tasks as more demanding, despite comparable objective metrics, pointed to the importance of subjective workload in understanding how age may impact driver experiences with automation. Future research should explore the interplay between subjective workload, fatigue, and contextual factors such as system reliability or environmental conditions, particularly over extended exposure to automation.

5.4.2 Effect of Task Conditions (Drive monitoring Vs NDRT) on Takeover Performance

This study highlighted the impact of Task Conditions on TOP across multiple metrics, corroborating existing research on driver performance in autonomous vehicles.

Both longitudinal as well as lateral SDLnP noted performance degradation under drive monitoring tasks, which is consistent with studies demonstrating that fatigue compromised lane-keeping performance, leading to increased variability in lateral and longitudinal control (May & Baldwin,

2009). However, this degradation in lane keeping was much greater under NDRT conditions. This could be understood from the Multiple Resource Theory framework (Wickens, 2002), where competing demands prior to engagement affected the primary task i.e. lane keeping (Weaver et al., 2020). However, it is also noteworthy that lane-keeping performance (lateral and longitudinal SDLnP) noted a higher variability due to the engagement in NDRT than of fatigue (Louw et al., 2019; Weaver, B. W. et al., 2020).

Interestingly, impairments in steering control were more pronounced under the drive monitoring condition than in the NDRT condition. Specifically, SWA variability and SWS were significantly higher during drive monitoring, suggesting greater difficulty in maintaining stable steering. The effect sizes indicate that the observed differences were substantial, with a large effect for SWA variability ($\eta^2 = 0.221$) and SWS ($\eta^2 = 0.167$). This finding aligns with the broader understanding that fatigue leads to a global decline in cognitive and motor functions (Abd-Elfattah, H. M., 2015). This decline in cognitive and motor functions affects the primary input after a TOR predominantly (Philip et al., 2005; Saxby et al., 2013; Niu S. et al., 2024). Likewise, SWS was more pronounced under drive monitoring, with a mean of 1.20 ($SD = 1.29$) compared to 0.55 ($SD = 0.55$) during NDRT. These results indicate diminished steering precision during fatigued states, consistent with findings that fatigue/drowsiness exacerbates steering variability and compromises vehicle control (Philip et al., 2005; Saxby et al., 2013). The large effect sizes for SDLnP and steering metrics suggest these impairments were robust across participants, emphasizing fatigue's pervasive impact on TOP.

Notably, task conditions did not significantly affect longitudinal deceleration or RTOT, contrasting with findings from Feldhütter et al. (2019) and Guo et al. (2021). However, lateral deceleration was significantly influenced by task conditions. While Jarosch et al. (2019) also reported no significant differences in longitudinal deceleration, this study diverged by finding a significant effect of task conditions on lateral deceleration. This discrepancy can be attributed to differences in task demands and road geometries. For instance, Du et al. (2015) found that fatigue impairs lane-keeping performance more severely on straight road segments, a condition consistent with the current study. Moreover, fatigue disproportionately affects tasks that require sustained attention and fine motor control, such as lateral adjustments and steering corrections (Niu S. et al., 2024). This study concurred with previous findings suggesting that decreased lateral control performance is a reliable indicator of both severe and moderate levels of fatigue. As van Loon et al. (2015) highlighted, fatigue influences the severity and timing of corrective steering behaviour. Fatigued drivers often employed more extreme steering angles to correct deviations further from their mean lateral position, indicating delayed and exaggerated responses to lane departures. Such corrective steering behaviours strongly signified fatigue and underscored the particular vulnerability of lateral control to fatigued states.

Another explanation for the increased variability in lateral control under drive monitoring conditions was that drivers (who were experiencing fatigue) may prioritize lateral stability when they perceived the driving task to be relatively easy. This aligned with findings from previous studies (Louw et al., 2015; Shahini et al., 2023; Zeeb et al., 2017), which suggested that drivers allocate more cognitive resources to lateral control when the perceived demand is low. In contrast, when engaged in NDRT, the resulting cognitive load disproportionately affected longitudinal performance, leading to greater inconsistencies in braking behaviour and deceleration (Louw et al., 2019; Zeeb et al., 2017).

The observation that subjective task demand perception remained consistent, even in fatigue-inducing scenarios, aligned with research indicating that fatigue and workload are distinct constructs governed by separate cognitive and physiological mechanisms. For instance, a study exploring the relationship among mental workload, occupational fatigue, physiological responses, and cognitive performance in office workers found a complex interplay among these factors, suggesting that mental workload and fatigue, while related, are distinct phenomena (Mahdavi et al., 2024).

5.4.3 Fatigue Induction and Mitigation Using Workload from NDRTs

5.4.3.1 *Fatigue Induction by Low Workload*

This study highlights the strong relationship between low workload and fatigue in autonomous vehicle contexts as evidenced by a significant effect on self-reported KSS levels, figure 5.10(a). Drive monitoring condition recorded higher levels of self-reported KSS ($M = 6.77$, $SD = 1.70$) than noted during the NDRT condition ($M = 3.50$, $SD = 1.53$), with the highest effect size measured $\eta^2 = 0.706$.

Figure 5.10 highlights a concerning trend: more than 80% of participants reported experiencing signs of drowsiness under drive monitoring conditions. Among them, nearly half (43.33%) struggled to stay awake, with some making considerable efforts to resist sleep. In contrast, only 10% described themselves as alert or rather alert, while another 10% reported feeling neither alert nor sleepy. These self-reported experiences were further validated through video observations, which revealed that 72% of participants exhibited visible signs of extreme fatigue – such as prolonged eye closures – between 18 and 22 minutes into the automated driving session. The consistency between subjective reports and objective video analysis underscores the rapid onset of fatigue and its potential risks for takeover performance in Level 3 automated vehicles. These findings are consistent with prior studies, which observed fatigue onset within 15–20 minutes of automation (Vogelpohl et al., 2019; Wu et al., 2019). Notably, in the study by Jarosch et al. (2019), no participants fell asleep, in contrast to three participants in Feldhütter et al. (2018) who fell asleep after 60 minutes of automated driving without task engagement, and eight participants in

Manabu et al. (2006). The significant effect of task conditions on KSS in this study supports the conclusion that low workload states are effective in inducing fatigue, even in simulated settings.

Interestingly, similar to Arefnezhad et al. (2022) age did not significantly affect fatigue levels, suggesting consistent effects across age groups. However, this contrasts with findings from Wu et al. (2019), which may stem from the small sample size of the current study. These results emphasize the universal challenge posed by fatigue in autonomous driving and underscore the necessity for strategies that address fatigue across diverse populations.

5.4.3.2 *Mitigating Fatigue Through NDRTs*

Engaging in NDRTs significantly mitigated fatigue, as evidenced by lower KSS scores in the NDRT condition. This supports the arousal hypothesis, which posits that cognitive engagement sustains alertness by activating neural mechanisms for attention and vigilance (May & Baldwin, 2009; Saxby et al., 2013). Enhanced alertness under NDRT conditions also improved lateral vehicle control and steering precision, with lower lateral SDLnP and reduced steering wheel angle variability, aligning with Song et al. (2017). However, engagement in NDRTs negatively impacted longitudinal control performance, leading to greater variability in longitudinal SDLnP and braking performance. Future research should investigate how to optimize the use of NDRTs in maintaining vigilance while not worsening performance at the same time.

RTOT was negligible, indicating that while NDRTs improve sustained control and situational awareness, they may not enhance rapid reactions to critical events. Future research should explore strategies to optimize NDRT design, balancing its role in sustaining vigilance without compromising overall driving performance, particularly in Level 3 automated vehicles.

5.5 Conclusion

This study underscores the critical influence of drowsiness and task conditions on TOP in autonomous vehicles. Fatigue, induced by low workload states, significantly impairs vehicle control, while engaging in NDRTs shows promise in mitigating these effects and enhancing situational awareness. However, the benefits of NDRTs may not extend to response timing, highlighting the need for further investigation into their design and implementation.

Age-related differences in objective TOP metrics were minimal, but “50+ age older drivers” reported higher perceived workloads, underscoring the importance of user-centric designs and tailored interventions to address subjective challenges.

These findings provide a foundation for developing safer and more adaptive autonomous vehicle systems. Future research should validate these results in real-world contexts and explore strategies to balance cognitive engagement with the demands of takeover scenarios to optimize driver performance and safety.

5.6 Limitations, Future Directions and Practical Implications

This study offers valuable insights but has several limitations. The use of a driving simulator may not fully capture the complexities of real-world autonomous driving; future studies should validate these findings in naturalistic settings. The small sample size, especially for age-related analyses, limits generalizability, necessitating larger, more diverse samples. Additionally, factors such as individual cognitive abilities and prior experience with autonomous vehicles, which may influence TOP, were not explored and warrant investigation. Finally, the non-significant effect of Task condition on RTOT suggests the need for experimental paradigms that incorporate varying urgency and complexity in takeover scenarios to better understand this relationship.

The findings have several practical implications for autonomous vehicle design. Mitigating fatigue should be a priority through engaging NDRTs or adaptive workload interventions, such as infotainment systems that promote alertness without causing distraction. The impact of NDRTs on specific TOP metrics highlights the need for task designs that enhance situational awareness and vehicle control without impairing response times. Moreover, the minimal effect of age on objective performance suggests that age-related differences may be less critical than perceived workload. Addressing these subjective challenges through user-centred designs and targeted training can improve the experience and safety of older drivers.

Chapter 6 Discussion

6.1 Summary of Previous Chapters

This thesis aimed to investigate and understand various factors – such as TB, NDRTs, fatigue, and age – that influence TOP in Level 3 automated vehicles, with a particular emphasis on the human factors involved in automation. To establish a strong foundation, the first chapter presented a systematic review of existing studies, mapping different elements of TORs to key human factors, including cognitive abilities, workload, and demographic influences such as age. The following sections summarize the key findings from each chapter.

In Chapter 2, a conceptual framework was introduced to map TOR to EF. This framework informed the design of a computer-based target detection experiment where each EF component had a corresponding task, which was – updating, shifting, inhibition, and combined. Findings from this experiment revealed that older adults exhibited generally slower responses, with a pronounced decline in the inhibition component of EF, leading to higher RT and error rates. Additionally, cognitive load and mental fatigue had a stronger impact on accuracy than on RT, suggesting that cognitive demands influence the correctness of responses more than their speed.

Chapter 3 examined the effect of TB (i.e. 2.5 s vs 5 s) and age on TOP, which was quantified into timeliness, as measured through RTOT, and quality, as measured through SWA, SWS, lateral SDLnP, longitudinal SDLnP, etc. Interestingly, age did not significantly influence TOP across different TB conditions. While older drivers generally drove at lower speeds than younger drivers, their overall takeover performance remained comparable. TB had a significant impact on steering behaviour, as reflected in changes in SWA and SWS, but it did not affect RTOT, SDLnP, or deceleration. Urgent TORs triggered reflexive and automatic responses, making RTOT differences non-significant. These findings suggested that TB primarily influenced immediate control inputs rather than broader stability metrics.

Chapter 4 investigated the impact of NDRTs on TOP with a TB of 5 seconds. Engagement in NDRTs significantly influenced SWA. The highest SWA variability was observed during the Observation task, indicating that monitoring the driving environment disrupted vehicle control the most. In contrast, the gaming task resulted in the lowest SWA variability, followed by video watching. Furthermore, engagement in NDRTs did not significantly affect RTOT, supporting previous research that suggests natural responses, such as regaining control of the vehicle, experience minimal interference from secondary task-related cognitive load.

Finally, chapter 5 focused on the effects of task condition (drive monitoring Vs engagement in NDRT) and age on TOP under a 3-second TB. Interestingly, age did not significantly affect

takeover performance across different drive monitoring conditions. While participants did not report differences in their perception of fatigue, they generally experienced higher workload in the drive monitoring conditions. Fatigue had a substantial impact on steering performance, while engagement in NDRTs had a more pronounced effect on SDLnP in both longitudinal and lateral dimensions. RTOT remained unaffected by fatigue, reinforcing previous findings that fatigue primarily degrades lateral control performance and may not necessarily affect the response time to an urgent stimulus. Additionally, lateral deceleration was significantly influenced by fatigue, supporting previous research that associates lateral instability with both moderate and severe fatigue levels in automated driving.

6.2 Methodological Considerations

The methodological framework of the thesis was designed to systematically investigate the factors influencing TOP in L3 AVs. By integrating various IVs and DVs while accounting for individual differences, this research adopted a comprehensive approach to studying takeover transitions.

A central focus of this thesis is understanding how different age group drivers respond to TORs in L3 AVs. The classification of older drivers as those aged 60 and above is grounded in both empirical research and practical considerations. The World Health Organization (WHO) and other international health and safety bodies commonly adopt this threshold due to observed declines in cognitive, sensory, and motor functions that significantly affect driving performance (WHO, 2015). Research indicates that age-related declines in executive functions – such as working memory, cognitive flexibility, and response inhibition – become more pronounced around this age, impacting situational awareness and response efficiency (Hasher & Zacks, 1988; Verhaeghen & Cerella, 2002). While certain ageing markers emerge as early as 40, they become substantially pronounced by 60, further reinforcing this classification (Lehallier et al., 2019; Shen et al., 2024).

Additionally, sensorimotor decline, including slower reaction times and reduced neuromuscular coordination, increasingly affects driving performance beyond 60, compromising the precision and stability of control inputs during takeovers (Seidler et al., 2010). From a regulatory standpoint, many jurisdictions impose stricter medical evaluations and driving assessments on drivers aged 60 and above to ensure road safety (Eby et al., 2015). Although older adults often possess extensive driving experience, their adaptability to automation may differ from younger drivers due to variations in technological familiarity and risk perception (Caird et al., 2005). Given the growing population of older drivers and their increasing interaction with automated vehicles, studying their takeover behaviour is crucial for informing policy decisions and optimizing in-vehicle support systems.

This thesis primarily employed a mixed-factorial design in multiple experiments, optimizing both within- and between-subject comparisons to enhance statistical power and ecological validity.

Task conditions were treated as repeated measures, allowing for within-subject comparisons across different workload manipulations such as TB, NDRT, and fatigue tasks. This approach enhanced statistical power while minimizing the required sample size. Age was included as between-subject factors to examine individual differences in takeover performance, see Figure 6.1.

The DVs focused on both the timeliness of takeover responses, measured through RTOT and the quality of responses, assessed through steering and lane-keeping metrics. TOP was assessed using multiple metrics to ensure a holistic evaluation of lateral and longitudinal control, see Figure 6.1. SWA and SWS captured the precision and stability of manual control inputs during takeovers, with a 2-degree SWA threshold marking takeover initiation. This threshold is related to the driver's first conscious intervention after a TOR, which is aligned with studies identifying steering correction as an early indicator of driver engagement (Gold et al., 2013). The threshold for steering and braking varies in different studies, some even do not explicitly define it, however,

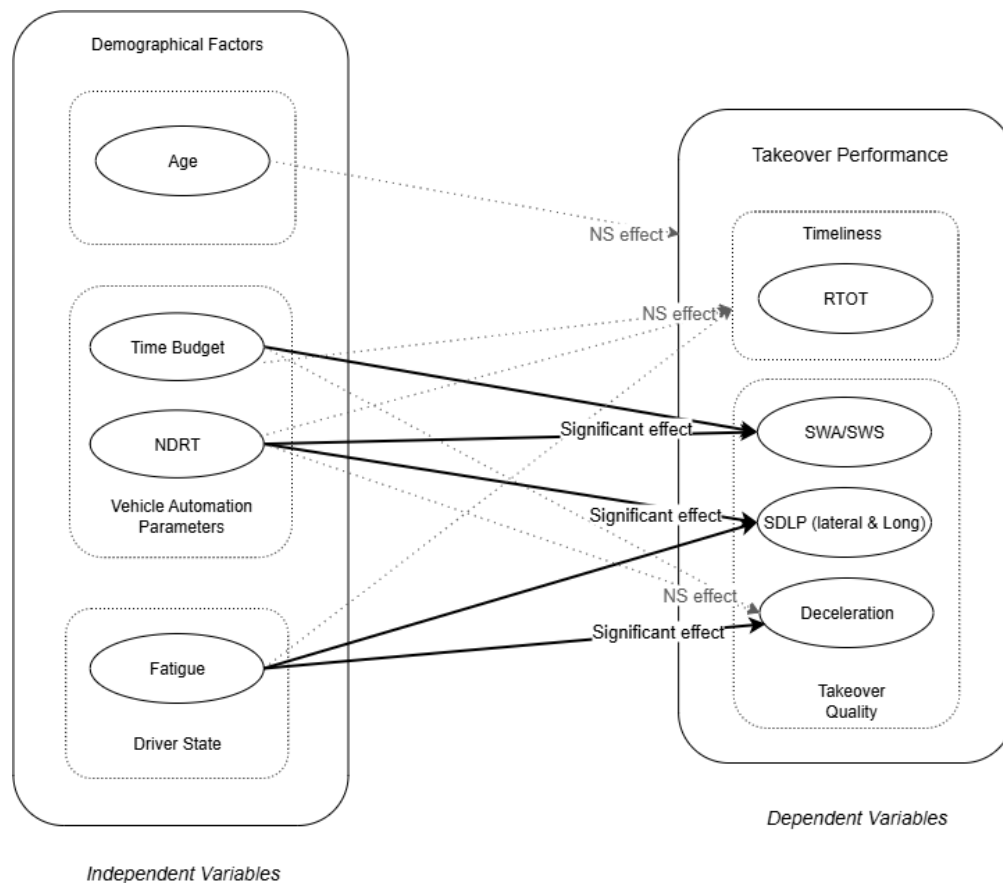


Figure 6. 1. Representing the factors affecting TOP and its quantified metrics.

the most frequently used are a 2° change of steering wheel angle and 10% actuation of the brake pedal (Cao Y. et al., 2021). Some have focused on reaction time to braking or accelerator input (Körber et al., 2016), offering a broader measure of engagement but potentially overlooking steering control intricacies. Other studies have suggested gaze reactions to be used as indicators of drivers' initial conscious efforts (Eriksson et al., 2020). Additionally, deceleration patterns were

analysed to evaluate controlled braking and its role in stabilization post-TOR. SDLnP was used to assess variability in lane-keeping and speed adjustments, with higher SDLnP values indicating impaired stabilization, particularly following prolonged automation (Zhang et al., 2023). By integrating SWA, SWS, SDLnP, and deceleration metrics, this thesis adopts a more comprehensive, index-based approach, capturing both the speed and stability of takeover responses.

This thesis employed self-reported workload measures such as NASA-TLX and the KSS. Future studies should include additional objective indices to enhance interpretability. Physiological measures, including heart rate variability and pupil dilation, offer real-time workload estimations. Combining objective workload indices with subjective ratings would yield a more nuanced understanding of cognitive demands during takeover scenarios.

Future research should refine age classification criteria, expand TOP metrics, and incorporate composite workload measures to further enhance the reliability and applicability of findings in automated driving contexts.

6.3 Empirical Contributions

This thesis provided valuable insights into the cognitive and behavioural mechanisms TOP in Level 3 automated vehicles, with important implications for the design of future autonomous driving systems.

6.3.1 Effect of Age on TOP

Ageing is associated with cognitive and motor function decline, which has traditionally raised concerns about the driving abilities of older adults. EF components such as updating, shifting, and inhibition tend to deteriorate with age, leading to slower reaction times and increased susceptibility to errors (Salthouse et al., 1998; Idowu et al., 2023). Moreover, older drivers often exhibit a decline in processing speed, affecting their ability to respond to dynamic driving scenarios (Salthouse et al., 1998; Medic-Pericevic S. et al., 2020). However, ageing is not solely characterized by declines in performance; many older adults develop compensatory strategies such as increased cautiousness, reduced driving speeds, and reliance on experience to maintain safe driving performance (Baltes & Baltes, 1990). In the context of L3 AVs, where drivers must regain control after a TOR, understanding these age-related changes becomes crucial. Given the intermittent nature of human intervention in L3 AVs, it is essential to investigate whether these compensatory strategies help mitigate potential age-related deficits in TOP.

To explore these issues, a series of controlled experiments were conducted in a high-fidelity driving simulator. The studies investigated how age, urgency, and cognitive load influenced takeover performance in L3 AVs. Participants were subjected to various TOR conditions that

differed in available TB and engagement in NDRTs. Measures such as RTOT, SWA, SDLnP, and deceleration were recorded to assess takeover timeliness and quality. Additionally, subjective workload measures were collected to examine cognitive demands across different age groups.

The findings across all chapters suggested that age did not significantly impact TOP under the conditions studied. Older adults generally exhibited slower reaction times in lab-based EF tasks, with a pronounced age-related decline in the inhibition component of EF, leading to higher RT and error rates (Idowu et al. 2023; Salthouse et al., 1998). This suggested that age-related declines may not be uniform across all EF components (Idowu et al. 2023). However, despite these declines, their overall ability to regain control in L3 AVs remained comparable to younger drivers across different TOR conditions, both in terms of takeover timeliness (Clark et al., 2017; Körber et al., 2016) and takeover quality (Favarò et al., 2019; Körber et al., 2016; Miller et al., 2016). The urgent nature of the TOR may have contributed to this comparability, as both older and younger drivers exhibited similar RTOT (Wu et al., 2022).

In terms of TOP quality, TB did not disproportionately affect older drivers. Although they usually drove at lower speeds, which served as a compensatory mechanism, their takeover quality – measured through SWA, SDLnP, deceleration, etc. – remained comparable to that of younger drivers as well. Similarly, when examining the impact of fatigue, older drivers reported higher subjective workload; however, this did not lead to a significant decline in takeover performance, contradicting assumptions that they are more vulnerable to automation-induced fatigue.

These findings suggest that chronological age alone may not be a decisive factor in takeover performance and that contextual factors such as urgency, cognitive load, and driving strategies have a more substantial impact on the ability to safely regain control in L3 AVs. The absence of significant age-related effects across these experiments suggests chronological age may not be an independent or exclusive factor affecting driving performance, rather driving context (urgent TORs), cognitive load, fatigue, etc., seems to have a higher influence in affecting TOP.

6.3.2 Effect of Workload

Workload is a fundamental factor influencing driving performance, particularly in L3 AVs, where drivers must intermittently resume control. Unlike manual driving, L3 automation introduces variability in workload demands, as drivers shift between passive monitoring and active control. The interplay between workload-related factors such as TB, NDRT, and fatigue significantly shapes TOP by modulating cognitive and motor demands. Effective workload management is essential to ensure safe and efficient takeovers, necessitating a deeper understanding of how these factors interact with cognitive resources and compensatory mechanisms.

The theoretical underpinnings of workload effects on driving can be explained through Multiple Resource Theory (MRT) (Wickens, 2002), the Compensatory Control Model (CCM) (Hockey,

1997), and broader frameworks of human attention and motor control. According to MRT, cognitive resources are divided into separate pools based on modality (e.g., visual, auditory), stage (e.g., perception, response), and processing code (e.g., spatial, verbal). If two tasks draw on the same resource pool, interference occurs, leading to performance degradation. This explains why high-workload conditions, such as engaging in NDRTs while driving, can impair TOP due to overlapping cognitive demands. The CCM posits that when workload increases, individuals adjust their behaviour to maintain primary task performance, often at the cost of increased mental effort or erratic control inputs. These theories provide a framework for understanding how workload factors, including TB, NDRT, and fatigue, influence takeover dynamics in L3 AVs.

TB is a critical workload-related factor in takeover performance, as it dictates the duration available for a driver to assess and respond to a TOR. The effect of TB (e.g., 2.5 s vs. 5 s) on TOP was examined in Chapter 3, where urgent TORs were initiated. The results indicated that RTOT differences between the two TB conditions were non-significant, suggesting that the urgency of the scenario (e.g., a broken-down vehicle blocking the lane) constrained RTOT within a limited range regardless of TB duration.

The influence of TB was more evident in steering-related metrics. A shorter TB (2.5 s) resulted in more abrupt and aggressive steering adjustments, as drivers had less time to assess the situation before taking action. In contrast, a longer TB (5 s) allowed for smoother and more deliberate control inputs, reducing SWA variability and excessive SWS. However, broader stability measures like SDLnP were not significantly affected by TB, reinforcing the idea that drivers prioritize immediate corrective actions (steering) before achieving full lane stabilization (Engström et al., 2017). This aligned with Gold et al. (2013), who proposed that in urgent TORs, drivers prioritize immediate vehicle control – primarily through rapid steering corrections – before engaging in higher-order stabilization tasks such as lane keeping and deceleration. These findings suggested that while TB did not directly alter reaction times, it plays a crucial role in shaping the quality of control execution, influencing steering precision and overall manoeuvre stability.

Similar findings emerged in Chapters 4 and 5, where neither NDRT engagement nor drive monitoring significantly affected RTOT. The absence of an RTOT effect aligns with prior research suggesting that under urgent conditions, drivers exhibit reflexive responses independent of secondary task engagement or fatigue levels. However, workload conditions did influence takeover quality, with distinct effects observed between different workload manipulations. Notably, the fatigue/observation condition was associated with larger SWA but smaller deviations in SDLnP, whereas the NDRT/gaming condition produced smaller SWA but larger SDLnP deviations.

6.3.2.1 *Effect of Fatigue/Observation Condition on SWA and SDLnP*

Fatigue is a psychobiological state resulting from prolonged cognitive or physical exertion, characterized by sensations of tiredness, diminished energy, an increased perception of effort, and impaired cognitive function, all contributing to decreased performance (Marcora et al., 2009; Smith et al., 2018). Individuals experiencing fatigue often exhibit increased reaction times accompanied by sleepiness, further degrading performance. It is a critical factor influencing driving performance, particularly in the context of AVs where prolonged passive monitoring can lead to reduced vigilance (Williamson et al., 2011). In L3 AVs, drivers are required to re-engage in driving upon receiving a TOR. However, cognitive and motor impairments induced by fatigue can compromise their ability to execute a safe and efficient takeover (Schmidt et al., 2009). Previous studies have shown that fatigue affects reaction times, attention allocation, and neuromuscular coordination, all of which are essential for effective vehicle control (Philip et al., 2005; Saxby et al., 2013). Unlike active secondary task engagement, such as NDRTs, where cognitive resources are allocated elsewhere, fatigue results in a global decline in cognitive and motor functions. Understanding how fatigue or passive vigilance tasks impact TOP is essential to designing effective countermeasures to mitigate safety risks in AVs. The Compensatory Control Model (CCM) (Hockey, 1997) describes how individuals adjust their control strategies when faced with increased task demands or physiological limitations such as fatigue. According to the model, when cognitive or motor resources are compromised, individuals redistribute effort to maintain primary task performance at the expense of secondary tasks. Furthermore, CCM suggests that performance trade-offs emerge when fatigued drivers attempt to sustain critical control tasks.

In chapters 4 and 5, both observation and drive monitoring task conditions were designed as low-workload vigilance tasks. The observation task lasted for six minutes, whereas the drive monitoring task extended to 50 minutes to induce drowsiness. Despite this difference, their effects on takeover performance were similar when compared to high-workload tasks such as NDRT gaming tasks.

Fatigue impaired sustained attention and motor coordination, affecting the ability to make precise steering movements, resulting in erratic steering behaviour and delayed overcorrections (Niu S. et al., 2024). Unlike secondary task engagement (as in NDRT), where attention was actively diverted, fatigue creates an overall degradation in performance due to cognitive and motor slowing. Such degradation in performance can have severe safety implications, particularly in AV takeovers where swift and precise motor responses are required. Other studies have also shown that fatigued drivers exhibited larger and more erratic steering corrections (Philip et al., 2005; Saxby et al., 2013). Our findings also aligned with van Loon et al. (2015), where fatigue

significantly affected both SWA and lateral lane deviation, but the degree to which SWA was affected was much higher.

As predicted by the Compensatory Control Model (Hockey, 1997), fatigued drivers unconsciously prioritize primary tasks (i.e. lane-keeping), resulting in lower SDLnP variability compared to NDRT at the cost of their steering inputs become erratic (larger, unpredictable SWA) due to impaired fine motor control (Niu S. et al., 2024). The erratic steering behaviour in fatigue was a direct result of degraded neuromuscular coordination, leading to delayed overcorrections rather than a deliberate reduction in control effort (as seen in NDRT). The strong impact of drive monitoring on SWA supported the idea that fatigue primarily impaired motor coordination rather than the lane-keeping strategy itself, reinforcing why lane variability (SDLnP) remains more stable compared to the extreme fluctuations observed in steering adjustments.

6.3.2.2 *Effects of NDRT on SWA and SDLnP*

As AVs evolve, drivers are increasingly able to engage in NDRTs during periods of automation. While this enhances comfort and usability, it also raises concerns regarding TOP, particularly in Level 3 automation, where drivers must resume control when prompted. Investigating the effects of NDRT engagement is essential to understanding how cognitive demands and attentional shifts impact driving performance. The ability to effectively transition from passive engagement in an NDRT to active vehicle control is a key safety consideration, as improper workload management can lead to unstable vehicle handling and impaired situational awareness.

When drivers engage in an NDRT, their visual and cognitive attention is diverted from driving, creating potential performance deficits upon resumption of control. According to Multiple Resource Theory (Wickens, 2002), tasks that compete for the same cognitive resources interfere with one another, leading to temporary impairments in performance. This interference is particularly pronounced in tasks that require EF rather than automated cognitive processes, as proposed by the Cognitive Control Hypothesis (Engström et al., 2017). Since lane-keeping is a controlled process that requires continuous attentional monitoring, engaging in an NDRT before a takeover increases lane variability, reflected in the higher SDLnP.

Research has consistently demonstrated that engagement in NDRTs leads to increased lateral lane variability, as drivers struggle to maintain precise lane position after resuming control (Zhang et al., 2023; Liang et al., 2022). Similarly, a meta-analysis by Weaver et al. (2020) found that engaging in NDRTs had a significant effect on lane-keeping performance, with a large effect size (Cohen's $d = 0.55$). This highlights that drivers engaged in NDRTs exhibit greater SDLnP compared to those who remain attentive. However, SWA was found to be minimally affected by NDRT engagement (Cohen's $d = 0.06$), suggesting that while lane-keeping deteriorates, drivers are still able to execute corrective steering inputs with relative precision.

These results underscored the importance of management of workload in automated vehicles, as both underload (leading to disengagement) and overload (leading to impaired control) can negatively impact takeover quality. A particularly noteworthy finding is that RTOT remained unaffected across conditions. This suggested that in urgent takeover scenarios, response initiation is largely automatic, supporting theoretical models that differentiate between controlled and reflexive responses to stimuli.

Taken together, these findings challenge traditional assumptions that reaction time alone determines takeover success. Instead, they emphasize the importance of cognitive flexibility, control precision, and workload management. Effective takeover transitions require not just quick reactions but also the ability to execute smooth and stable control inputs.

6.4 Theoretical contributions

The findings from this thesis contribute significantly to validating and improving the existing theoretical frameworks in the domain of TOP in L3 AVs. By examining the effects of workload-related factors such as TB, NDRTs, and fatigue, this research provided novel insights into the cognitive and motor mechanisms underpinning takeover transitions. These results allowed for a critical evaluation of established theories, including Multiple Resource Theory, the Cognitive Control Hypothesis (CCH), and the Compensatory Control Model (CCM) leading to a refined theoretical landscape for understanding takeover performance in AV contexts.

6.4.1 Multiple Resource Theory (MRT), Cognitive Control Hypothesis (CCH) and Takeover Performance

MRT (Wickens, 2002) postulates that cognitive resources are divided into multiple pools based on modality, processing stage, and task code. Interference arises when two tasks draw on the same resource pool, leading to degraded performance. The CCH (Engström et al., 2017) differentiates between controlled and automatic processes in driving. Controlled processes, governed by EF, are vulnerable to interference from NDRTs, whereas automatic processes remain stable.

Findings from this thesis aligned with MRT and CCH by demonstrating that NDRT engagement impaired lane-keeping (higher SDLnP) due to shared cognitive resource demands. When drivers are engaged in NDRTs, their ability to allocate attention to lateral control is diminished, leading to increased lane deviations. However, RTOT remained unaffected, suggesting that primary control initiation operated through a separate, more reflexive mechanism. This calls for a refinement of MRT, emphasizing that while resource competition affected controlled processes (e.g., lane-keeping), reflexive processes and automated processes (e.g., initial steering corrections) may remain intact in urgent scenarios, which is in line with the CCH hypothesis.

6.4.2 Compensatory Control Model and Takeover Performance

The CCM (Hockey, 1997) suggests that individuals adapt their control strategies to compensate for increased workload by prioritizing essential tasks. This study found that fatigued drivers prioritized lane-keeping (lower SDLnP variability) at the cost of erratic steering corrections (higher SWA deviations), which aligns with the CCM's prediction that performance trade-offs emerge under different workload conditions. While the CCM accounts for compensatory strategies, this study highlighted that compensatory behaviours in automation scenarios are more nuanced than previously thought. Instead of deliberate compensation, the erratic SWA observed in fatigued drivers likely reflected degraded neuromuscular coordination along with unconscious prioritization of lane-keeping. This suggests that compensatory control under fatigue is not a purely cognitive adaptation but also involves involuntary motor impairments, necessitating an integration of motor control theories into CCM for automated driving contexts.

6.4.3 Proposed Framework

These findings suggested that a dual-process framework may better explain takeover performance. While the CCH correctly identifies EF-driven impairments, the sharp increase in SWA under the influence of fatigue implies an additional degradation of lower-level motor coordination mechanisms. This necessitates a more comprehensive model integrating both cognitive control and neuromuscular degradation effects, particularly for prolonged automation exposure.

Based on these considerations, this thesis proposed the Takeover Adaptation Model (TAM), which integrates elements of MRT, CCH, and CCM into a unified framework. TAM postulates that takeover performance is governed by three interacting dimensions:

- Cognitive resource competition and interference effects result from task demands, aligning with MRT.
- Tasks that require EF control experience a greater level of interference from a secondary cognitive task, while the task requiring only automatic responses may observe minimal or lesser effects, extending the CCH framework.
- Underload and fatigue lead to motor degradation, prompting compensatory adjustments and prioritization of critical tasks, as described by CCM.

Figure 6.2 illustrates the relationship between workload levels, driving performance, and associated theoretical models. The x-axis represents workload level, ranging from underload (left) to overload (right), while the y-axis denotes performance. The figure follows an inverted U-shaped trend, indicating that performance is optimal at moderate workload levels but deteriorates under extreme underload or overload conditions, supporting the Yerkes-Dodson Law.

The left side of the figure is grounded in the Compensatory Control Model, which suggests that under low workload conditions, such as those arising from vigilance tasks, drivers may experience underload and fatigue. The state of underload/fatigue is associated with loss of fine motor control and precision, overcorrection of steering behaviour, and prioritization of lane-keeping over steering performance, per CCM. Conversely, the right side of the figure is based on the Multiple Resource Theory and Cognitive Control Hypothesis, which argue that high cognitive workload – such as from engaging in an NDRT like gaming – can lead to overload and stress. In this state, CL from the NDRT interferes with driving performance (i.e. lane keeping performance), particularly when the subtask is a controlled process. Nonetheless, the initial transition cost when switching from the NDRT to driving is low because the driver is already engaged in an activity, resulting in minimal effects on early SWA. Whereas in a fatigued or underloaded state, the lack of engagement causes greater disruption, resulting in a larger SWA. However, lane-keeping performance, as a controlled process, is significantly affected in overload conditions.

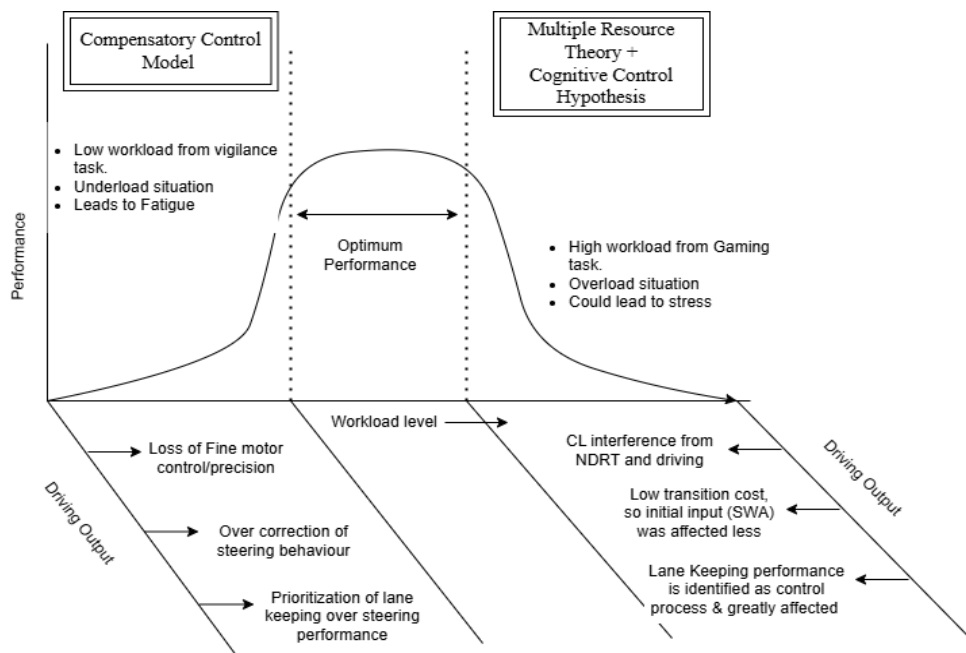


Figure 6. 2. Conceptual Framework detailing the effect of changes in workload on performance and driving output.

Overall, the model supports the notion that moderate workload levels facilitate optimal driving performance, while both excessive and insufficient workload impair key driving behaviours. This aligned with prior research on workload-performance trade-offs in automated driving contexts (Saxby et al. 2013; Miller et al. 2015; Körber et al. 2015; Gonçalves et al. 2016; Weaver et al., 2022). It provides a more holistic perspective by recognizing that takeover performance was not solely dictated by cognitive workload but also by the interplay between cognitive flexibility, motor stability, and adaptive compensation.

6.5 Conclusion and Limitations

This thesis systematically investigated the factors influencing TOP in L3 AVs, focusing on age-related differences, workload variations, and their interplay with takeover timeliness and quality. Contrary to common assumptions, older drivers demonstrated comparable TOP to their younger counterparts. While previous laboratory-based studies have consistently reported age-related slowing in RT, our findings suggest that this decline does not directly translate into delayed response times in takeover scenarios. Older drivers' RTOT remained on par with younger drivers, and their takeover quality – measured through steering and lane-keeping stability – did not significantly differ. Several factors may explain this outcome, including the urgent nature of the TORs used in the study, the compensatory strategies employed by older drivers (such as cautious driving and lower speeds), and the possibility that the sample comprised high-functioning older adults with preserved cognitive and motor capabilities. Additionally, extensive driving experience and risk-averse behaviours may have provided older drivers with the ability to respond effectively, despite potential declines in executive functioning.

Beyond age-related effects, this research also examined how workload – modulated by TB, NDRTs, and fatigue – affects takeover performance. The results revealed that workload influences TOP in distinct ways depending on the context. Fatigue, resulting from prolonged automation with low cognitive engagement, had a pronounced impact on initial steering corrections (SWA), whereas lane-keeping stability (SDLnP) remained relatively unaffected. Conversely, in high-workload conditions induced by engaging NDRTs, primary steering inputs showed reduced deviations, but lane-keeping performance deteriorated. These findings suggest that workload does not uniformly impair driving performance; rather, its effects manifest differently based on whether the primary impairment is in motor coordination (as in fatigue) or attentional control (as in NDRT engagement).

While this study provides valuable insights, several methodological considerations should be addressed in future research. First, the use of a driving simulator, limited choice of NDRT tasks, etc., while advantageous for experimental control, may not fully replicate the complexities of real-world driving. Additionally, due to software limitations, key longitudinal metrics – such as Time-to-Collision and headway distance – could not be calculated. The inclusion of these measures would have provided deeper insights into driver response and takeover dynamics. Future studies should validate these findings in naturalistic settings to assess the generalizability of these effects. Second, the small sample size, particularly for age-related analyses, limits statistical power and warrants further research with larger and more demographically diverse populations. Third, individual cognitive abilities, neuromuscular health, and prior exposure to automated driving technologies were not explicitly accounted for in this study but could play a crucial role in

determining takeover performance. Future research should explore how these factors interact with age and workload to refine our understanding of driver readiness and adaptability in L3 AVs.

Overall, this thesis contributes to the growing body of research on human-automation interaction by challenging simplistic assumptions about ageing and takeover performance, highlighting the complex role of workload in driving transitions and advocating for a shift toward more adaptive, context-aware automation strategies. By bridging theoretical insights with practical design considerations, this research provides a foundation for developing safer and more intuitive automated driving systems that account for the diverse needs of human drivers.

6.6 Future Research

While this thesis has provided critical insights into TOP in L3 AVs, several open questions remain. A deeper understanding of individual differences in takeover readiness, the role of cognitive flexibility, and the potential for adaptive automation systems could significantly enhance the safety and efficiency of future AVs. Building upon these findings, a future (postdoctoral) research agenda could focus on developing and validating adaptive driver monitoring and intervention strategies for optimizing takeover performance in L3 AVs.

The primary aim of this research would be to design and validate a multi-modal driver state monitoring system that dynamically adjusts automation parameters – such as time budget and alerting strategies – based on real-time assessments of driver readiness, cognitive workload, and fatigue. This research will build on prior work and address key limitations by integrating physiological, behavioural, and performance-based measures into a holistic framework for takeover optimization.

A key question this research seeks to answer is how individual differences in cognitive flexibility, driving experience, and fatigue resilience interact with automation demands to influence takeover stability. While previous studies have largely treated age as a categorical variable, this research will refine the approach by considering drivers as a heterogeneous group, classified based on cognitive and behavioural profiles rather than chronological age alone. Additionally, the study will explore how adaptive strategies, such as dynamic TOR timing and modality selection, can enhance takeover success by aligning intervention strategies with driver states and situational demands. Future research will also propose an objective measure to properly define “takeover success”.

6.6.1 Proposed Methodology

The research will adopt a multi-phase experimental approach, combining driving simulation studies and real-world validation to develop and refine an adaptive automation framework.

The first phase will focus on understanding individual differences in takeover readiness, particularly the role of cognitive flexibility and driving experience across a broader age range. A driving simulator study will assess cognitive function, fatigue resilience, and adaptability to automation, collecting physiological (e.g., heart rate variability, eye-tracking) and performance-based (e.g., reaction variability, lane-keeping stability) indicators. This phase aims to develop a predictive model linking individual traits to takeover performance.

The second phase will build on these findings by developing and testing adaptive strategies that modify takeover parameters based on real-time driver states. This includes dynamically adjusting the time budget based on workload levels and selecting the most effective TOR modality (e.g., auditory, haptic, visual) depending on situational urgency and automation duration. Controlled experiments will evaluate the impact of these adaptive strategies on takeover performance, measuring stability and precision in steering and braking responses.

The final phase will integrate these adaptive components into a framework for personalized automation. This framework will move beyond static takeover interventions by incorporating machine learning models that continuously adapt automation behaviour to the individual driver. By tailoring system responses based on user profiles and real-time monitoring, this research aims to create a more intuitive and supportive automation experience that maximizes safety while maintaining driver trust.

6.6.2 Expected Contributions and Impact

This research will extend existing models of human-automation interaction by incorporating real-time cognitive state assessments into adaptive automation strategies. Methodologically, it will introduce a multi-modal, index-based approach to evaluating takeover readiness beyond traditional reaction time measures. By developing a personalized adaptive takeover system, the study will contribute to the advancement of intelligent in-vehicle support systems, improving both the usability and safety of L3 AVs.

6.7 Key Findings

By examining individual differences, the impact of workload factors such as TB, NDRT, and fatigue, and by refining takeover performance measurement, this research advances our understanding of driver readiness and automation design.

6.7.1 Empirical Findings

- **Older Adults and Takeover Performance:** Contrary to expectations, older adults (60+) demonstrated takeover timeliness (RTOT) and takeover quality comparable to younger drivers. While cognitive ageing resulted in slowed reaction times in laboratory settings,

the urgency of TORs and older adults' compensatory strategies (e.g., cautious driving) mitigated age-related decline.

- The relationship between workload and takeover quality followed a non-linear pattern. Underload (leading to fatigue) had a pronounced effect on initial steering inputs, whereas higher engagement conditions primarily affected lane-keeping precision (SDLnP). This underscores the importance of maintaining optimal cognitive engagement during automation.
- Engagement in NDRTs impaired lateral stability (SDLnP) but did not significantly impact RTOT. This suggests that while secondary tasks can affect takeover precision, the initiation of response remained relatively unaffected in urgent situations.
- Fatigue and Automation: Fatigue negatively influenced lane-keeping and control precision, with significant deviations in SWA and SDLnP emerging within 15-20 minutes of automated driving. These findings reinforced the need for adaptive countermeasures to mitigate fatigue-related impairments in automated driving.
- Reaction time alone was not predictive of overall takeover quality. Instead, successful transitions depended on cognitive flexibility and controlled execution of lane-keeping and steering adjustments.

6.7.2 Theoretical Contributions

- This research extends Multiple Resource Theory by demonstrating that takeover performance is governed by both controlled (lane-keeping adjustments) and reflexive (steering corrections) processes.
- This study challenged the assumption that chronological age alone determines automation readiness. Instead, individual differences in cognitive flexibility, driving experience, and risk perception play a more significant role.
- Findings suggested that workload-induced variations in takeover performance are context-dependent. Underload leads to fatigue-driven impairments in early response mechanisms (steering), whereas overload disproportionately affects secondary responses (lane-keeping).
- The distinction between reaction time and takeover quality highlights the necessity of shifting focus from mere response speed to the controlled execution of takeover manoeuvres.

References

- 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. <https://doi.org/https://doi.org/10.1016/j.aap.2011.03.023>
- 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.
- Arca, A. A., Chouljian, E., & Mouloua, M. (2022). The role of gender differences in distracted driving behavior: a psychophysiological approach. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
- Arefnezhad, S., Eichberger, A., & Koglbauer, I. V. (2022). Effects of automation and fatigue on drivers from various age groups. *Safety*, 8(2), 30.
- Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and physiological optics*, 34(5), 502-508.
- Ashendorf, L., & McCaffrey, R. J. (2008). Exploring age-related decline on the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist*, 22(2), 262-272.
- Barra, J., Bray, A., Sahni, V., Golding, J. F., & Gresty, M. A. (2006). Increasing cognitive load with increasing balance challenge: recipe for catastrophe. *Exp Brain Res*, 174(4), 734-745. <https://doi.org/10.1007/s00221-006-0519-2>
- Behrens, M., Gube, M., Chaabene, H., Prieske, O., Zenon, A., Broscheid, K.-C., Schega, L., Husmann, F., & Weippert, M. (2023). Fatigue and human performance: an updated framework. *Sports medicine*, 53(1), 7-31.
- Beukel, A. P. v. d., & Voort, M. C. v. d. (2013, 6-9 Oct. 2013). The influence of time-criticality on Situation Awareness when retrieving human control after automated driving. 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013),
- Boksem, M. A., & Tops, M. (2008). Mental fatigue: costs and benefits. *Brain research reviews*, 59(1), 125-139.
- Bolton, D. A., & Staines, W. R. (2012). Age-related loss in attention-based modulation of tactile stimuli at early stages of somatosensory processing. *Neuropsychologia*, 50(7), 1502-1513.

Brandenburg, S., & Backhaus, N. (2016). The dynamics of film-induced affect and its effect on the interaction with tablet PCs. *Behaviour & Information Technology*, 35(5), 410-421.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, 39(4), 713–726. [https://doi.org/10.1016/S0896-6273\(03\)00466-5](https://doi.org/10.1016/S0896-6273(03)00466-5)

Budescu, D. V. (1982). The Power of the F Test in Normal Populations with Heterogeneous Variances. *Educational and Psychological Measurement*, 42(2), 409-416. <https://doi.org/10.1177/001316448204200202>

Caird, J. K., Edwards, C. J., Creaser, J. I., & Horrey, W. J. (2005). Older driver failures of attention at intersections: Using change blindness methods to assess turn decision accuracy. *Human Factors*, 47(2), 235-249.

Caldwell, J. A. (2009, March). Methodological issues associated with measuring the operational impact of fatigue in aviation. In *Proceedings of the International Conference on Fatigue Management in Transport Operations*.

Cao, Y., Zhou, F., Pulver, E. M., Molnar, L. J., Robert, L. P., Tilbury, D. M., & Yang, X. J. (2021, September). Towards standardized metrics for measuring takeover performance in conditionally automated driving: A systematic review. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 65, No. 1, pp. 1065-1069). Sage CA: Los Angeles, CA: SAGE Publications.

Carr, D. B., Duchek, J. M., Meuser, T. M., & Morris, J. C. (2006). Older adult drivers with cognitive impairment. *American family physician*, 73(6), 1029-1034.

Chao, H. F., Chen, M. S., & Kuo, C. Y. (2022). Attention modulates the contextual similarity effect in negative priming: evidence from task demand and attentional capture. *Memory*, 30(7), 895-914. <https://doi.org/10.1080/09658211.2022.2058017>

Chihuri, S., Mielenz, T. J., DiMaggio, C. J., Betz, M. E., DiGuseppi, C., Jones, V. C., & Li, G. (2016). Driving cessation and health outcomes in older adults. *Journal of the American Geriatrics Society*, 64(2), 332-341.

Choudhary, P., & Velaga, N. R. (2017). Modelling driver distraction effects due to mobile phone use on reaction time. *Transportation research part C: emerging technologies*, 77, 351-365.

Cicchino, J. B. (2015). Why have fatality rates among older drivers declined? The relative contributions of changes in survivability and crash involvement. *Accident Analysis & Prevention*, 83, 67-73.

Clark, H., & Feng, J. (2017). Age differences in the takeover of vehicle control and engagement in non-driving-related activities in simulated driving with conditional automation. *Accident Analysis & Prevention*, 106, 468-479.

Cooper, J. M., Wheatley, C. L., McCarty, M. M., Motzkus, C. J., Lopes, C. L., Erickson, G. G., Baucom, B. R., Horrey, W. J., & Strayer, D. L. (2020). Age-related differences in the cognitive, visual, and temporal demands of in-vehicle information systems. *Frontiers in psychology*, 11, 1154.

Coyne, R., Ryan, L., Moustafa, M., Smeaton, A. F., Corcoran, P., & Walsh, J. C. (2023). Assessing the physiological effect of non-driving-related task performance and task modality in conditionally automated driving systems: A systematic review and meta-analysis. *Accident Analysis & Prevention*, 192, 107243.

Danckert, J., Hammerschmidt, T., Marty-Dugas, J., & Smilek, D. (2018). Boredom: Under-aroused and restless. *Consciousness and cognition*, 61, 24-37.

De Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27, 196-217.

Deary, I. J., Johnson, W., & Starr, J. M. (2010). Are processing speed tasks biomarkers of cognitive aging? *Psychology and aging*, 25(1), 219.

Department, o. T. (2022, November 24). Reported road casualties in Great Britain: younger driver factsheet, 2021. Retrieved from <https://www.gov.uk/government/statistics/reported-road-casualties-great-britain-older-and-younger-driver-factsheets-2021/reported-road-casualties-in-great-britain-younger-driver-factsheet-2021>

Desmond, P. A., & Hancock, P. A. (2000). Active and passive fatigue states. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 455–465). CRC Press.

Detjen, H., Pfleging, B., & Schneegass, S. (2020). A wizard of oz field study to understand non-driving-related activities, trust, and acceptance of automated vehicles. Paper presented at the 12th International conference on automotive user interfaces and interactive vehicular applications.

- Dogan, E., Honnêt, V., Masfrand, S., & Guillaume, A. (2019). Effects of non-driving-related tasks on takeover performance in different takeover situations in conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, 494-504.
- Dogan, E., Yousfi, E., Bellet, T., Tijus, C., & Guillaume, A. (2021). Manual takeover after highly automated driving: Influence of budget time and lane change assist on takeover performance. *Proceedings of the 32nd European Conference on Cognitive Ergonomics*,
- Doubek, F., Loosveld, E., Happee, R., & De Winter, J. (2020). Takeover Quality: Assessing the Effects of Time Budget and Traffic Density with the Help of a Trajectory-Planning Method. *Journal of advanced transportation*, 2020(1), 6173150.
- Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: An event-related fMRI study. *Cognitive Brain Research*, 9(1), 103–109.
- Du, H., Zhao, X., Zhang, X., Zhang, Y., & Rong, J. (2015). Effects of fatigue on driving performance under different roadway geometries: a simulator study. *Traffic injury prevention*, 16(5), 468-473.
- Du, N., Zhou, F., Pulver, E. M., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Predicting driver takeover performance in conditionally automated driving. *Accident Analysis & Prevention*, 148, 105748.
- Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The unengaged mind: Defining boredom in terms of attention. *Perspectives on psychological science*, 7(5), 482-495.
- Eby, D. W., Molnar, L. J., Zhang, L., St Louis, R. M., Zanier, N., & Kostyniuk, L. P. (2015). Keeping older adults driving safely: A research synthesis of advanced in-vehicle technologies, A LongROAD study.
- Endsley, M. R. (1997). Situation awareness, automation and free flight. *USA/Europe ATM R&D Seminar*,
- Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, 42(3), 462-492. doi:10.1080/001401399185595

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of Cognitive Load on Driving Performance: The Cognitive Control Hypothesis. *Human Factors*, 59(5), 734-764. doi:10.1177/0018720817690639

Eriksson, A., Petermeijer, S. M., Zimmermann, M., De Winter, J. C., Bengler, K. J., & Stanton, N. A. (2018). Rolling out the red (and green) carpet: supporting driver decision making in automation-to-manual transitions. *IEEE Transactions on Human-Machine Systems*, 49(1), 20-31.

Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: noncritical transitions to and from manual control. *Human Factors*, 59(4), 689-705.

Eriksson, A., Banks, V., & Stanton, N. (2017). Transition to manual: Comparing simulator with on-road control transitions. *Accident Analysis & Prevention*, 102, 227-234.

Faber, L. G., Maurits, N. M., & Lorist, M. M. (2012). Mental fatigue affects visual selective attention. *PLoS ONE*, 7(10), e48073. <https://doi.org/10.1371/journal.pone.0048073>

Faber, K., & van Lierop, D. (2020). How will older adults use automated vehicles? Assessing the role of AVs in overcoming perceived mobility barriers. *Transportation research part A: policy and practice*, 133, 353-363.

Fatima, K., Moridpour, S., De Gruyter, C., & Saghapour, T. (2020). Elderly Sustainable Mobility: Scientific Paper Review. *Sustainability* (Switzerland), 12(18). <https://doi.org/10.3390/su12187319>

Favarò, F. M., Seewald, P., Scholtes, M., & Eurich, S. (2019). Quality of control takeover following disengagements in semi-automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 196-212.

Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017). How the duration of automated driving influences take-over performance and gaze behavior. *Advances in ergonomic design of systems, products and processes: Proceedings of the annual meeting of GfA 2016*,

Feldhütter, A., Kroll, D., & Bengler, K. (2018). Wake up and take over! The effect of fatigue on the take-over performance in conditionally automated driving. *2018 21st international conference on intelligent transportation systems (itsc)*,

Feldhütter, A., Ruhl, A., Feierle, A., & Bengler, K. (2019). The effect of fatigue on take-over performance in urgent situations in conditionally automated driving. *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*,

- Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017). How the duration of automated driving influences take-over performance and gaze behavior. Paper presented at the Advances in ergonomic design of systems, products and processes: Proceedings of the annual meeting of GfA 2016.
- Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86, 186-204.
- Gall, B. (2006). A guide to improving alertness through effective management of fatigue. *Petroleum Review*, 60(714), 24–25.
- Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). “Take over!” How long does it take to get the driver back into the loop? Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57(1), 1938-1942. doi:10.1177/1541931213571433
- Gold, C., & Bengler, K. (2014). Taking over control from highly automated vehicles. *Advances in Human Aspects of Transportation: Part II*, 8, 64.
- Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). “Take over!” How long does it take to get the driver back into the loop? Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57(1), 1938-1942. <https://doi.org/10.1177/1541931213571433>
- Gonçalves, J., Happee, R., & Bengler, K. (2016). Drowsiness in conditional automation: proneness, diagnosis and driving performance effects. 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC),
- Gormley, M., & O'Neill, D. (2019). Driving as a Travel Option for Older Adults: Findings From the Irish Longitudinal Study on Aging [Original Research]. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01329>
- Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, 13(7), 491-505. <https://doi.org/10.1038/nrn3256>
- Guo, H., Zhang, Y., Cai, S., & Chen, X. (2021). Effects of Level 3 Automated Vehicle Drivers' Fatigue on Their Take-Over Behaviour: A Literature Review. *Journal of advanced transportation*, 2021(1), 8632685.
- Guo, H., Qiu, H., Zhou, Y., & Deng, Y. (2024). Analysis of Influencing Factors of Level 3 Automated Vehicle Takeover: A Literature Review. *Sustainability* (2071-1050), 16(19).

Gwyther, H., & Holland, C. (2012). The effect of age, gender and attitudes on self-regulation in driving. *Accident Analysis & Prevention*, 45, 19-28.

Hadi, A. M., Li, Q., Wang, W., Yuan, Q., & Cheng, B. (2020). Influence of passive fatigue and take-over request lead time on drivers' take-over performance. *Advances in Human Aspects of Transportation: Proceedings of the AHFE 2020 Virtual Conference on Human Aspects of Transportation*, July 16-20, 2020, USA,

Hakamies-Blomqvist, L. (1998). Why do older drivers give up driving. *Accid Anal Prev*, 30(3), 3015-3312.

Hakamies-Blomqvist, L., Siren, A., & Davidse, R. J. (2004). OLDER DRIVERS - A REVIEW.

Hale, S., Myerson, J., Smith, G. A., & Poon, L. W. (1988). Age, variability, and speed: between-subjects diversity. *Psychol Aging*, 3(4), 407-410. <https://doi.org/10.1037//0882-7974.3.4.407>

Hancock, P. A., Lesch, M., & Simmons, L. (2003). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis & Prevention*, 35(4), 501-514.

Hancock, P. A., Nourbakhsh, I., & Stewart, J. (2019). On the future of transportation in an era of automated and autonomous vehicles. *Proceedings of the National Academy of Sciences*, 116(16), 7684-7691.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Hartley, L. (2004). *Traffic and transport psychology: Theory and application*. Fatigue and driving, Elsevier, 221-229.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, 22, 193–225.

He, J., McCarley, J. S., & Kramer, A. F. (2014). Lane Keeping Under Cognitive Load: Performance Changes and Mechanisms. *Human Factors*, 56(2), 414-426. doi:10.1177/0018720813485978

Hecht, T., Feldhütter, A., Draeger, K., & Bengler, K. (2020). What do you do? An analysis of non-driving related activities during a 60 minutes conditionally automated highway drive. Paper presented at the Human Interaction and Emerging Technologies: Proceedings of the 1st International Conference on Human Interaction and Emerging Technologies (IHiet 2019), August 22-24, 2019, Nice, France.

- Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature Reviews Neuroscience*, 5(2), 87-96. <https://doi.org/10.1038/nrn1323>
- Hess, T. M., & Knight, R. C. (2021). Adult age differences in the effects of chronic mental fatigue on task-related fatigue, appraisals, and performance. *Motivation science*, 7(2), 122.
- Hohenberger, C., Spörrle, M., & Welp, I. M. (2016). How and why do men and women differ in their willingness to use automated cars? The influence of emotions across different age groups. *Transportation Research Part A: Policy and Practice*, 94, 374-385.
- Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020). The 'Real-World Approach' and Its Problems: A Critique of the Term Ecological Validity [Conceptual Analysis]. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00721>
- Hoonakker, P., Carayon, P., Gurses, A. P., Brown, R., Khunlertkit, A., McGuire, K., & Walker, J. M. (2011). Measuring workload of ICU nurses with a questionnaire survey: the NASA Task Load Index (TLX). *IIE transactions on healthcare systems engineering*, 1(2), 131-143.
- Hoshino, K., Suzuki, A., Honda, M., & Hara, T. (2023). Designing Auditory Warnings for Preventing Arousal-level Decline during Train Operation and Examining Its Effectiveness. *Quarterly Report of RTRI*, 64(4), 229-235.
- Huang, G., & Pitts, B. (2020). Age-related differences in takeover request modality preferences and attention allocation during semi-autonomous driving. In Q. Gao & J. Zhou (Eds.), 6th International Conference on Human Aspects of IT for the Aged Population, ITAP 2020, held as part of the 22nd International Conference on Human-Computer Interaction, HCII 2020 (Vol. 12207 LNCS, pp. 135-146): Springer.
- Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science*, 102, 1-13.
- Hultsch, D. F., MacDonald, S. W., & Dixon, R. A. (2002). Variability in reaction time performance of younger and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(2), P101-P115.
- IIHS, I. I. f. H. S. (2020). Older drivers. Retrieved 24 December 2024 from <https://www.iihs.org/topics/older-drivers>

Idowu, M. I., & Szameitat, A. J. (2023). Executive function abilities in cognitively healthy young and older adults-A cross-sectional study. *Front Aging Neurosci*, 15, 976915. <https://doi.org/10.3389/fnagi.2023.976915>

Jamson, A. H., Merat, N., Carsten, O. M., & Lai, F. C. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transportation research part C: emerging technologies*, 30, 116-125.

Jamson, A. H., Lai, F. C., & Carsten, O. M. (2008). Potential benefits of an adaptive forward collision warning system. *Transportation research part C: emerging technologies*, 16(4), 471-484.

Jarosch, O., Bellem, H., & Bengler, K. (2019). Effects of task-induced fatigue in prolonged conditional automated driving. *Human factors*, 61(7), 1186-1199.

Jeong, H., & Liu, Y. (2019). Effects of non-driving-related-task modality and road geometry on eye movements, lane-keeping performance, and workload while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, 157-171.

Jiang, K., Ling, F., Feng, Z., Wang, K., & Shao, C. (2017). Why do drivers continue driving while fatigued? An application of the theory of planned behaviour. *Transportation Research Part A: Policy and Practice*, 98, 141-149.

Jin, M., Lu, G., Chen, F., Shi, X., Tan, H., & Zhai, J. (2021). Modeling takeover behavior in level 3 automated driving via a structural equation model: Considering the mediating role of trust. *Accident Analysis & Prevention*, 157, 106156.

K, M. B. S. (Jan 16, 2023). Frequently Asked Questions About Autonomous Vehicles. W&B Fully Connected. Retrieved 26 Novemeber 2023, from <https://wandb.ai/madhana/faq-autonomous-vehicles/reports/Frequently-Asked-Questions-About-Autonomous-Vehicles--VmldzoyNTgwMzA0>

Kausler, D. H. (1991). *Experimental psychology, cognition, and human aging*, 2nd ed [doi:10.1007/978-1-4613-9695-6]. Springer-Verlag Publishing. <https://doi.org/10.1007/978-1-4613-9695-6>

Kenemans, J. L., Smulders, F. T., & Kok, A. (1995). Selective processing of two-dimensional visual stimuli in young and old subjects: Electrophysiological analysis. *Psychophysiology*, 32(2), 108-120.

- Khan, B. A., Leva, M. C., & Cromie, S. (2021). A Systematic Review of Older Drivers in a Level 3 Autonomous Vehicle: A Cognitive Load Perspective. *Human Mental Workload: Models and Applications: 5th International Symposium, H-WORKLOAD 2021, Virtual Event, November 24–26, 2021, Proceedings 5*,
- Kluger, B. M., Krupp, L. B., & Enoka, R. M. (2013). Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. *Neurology*, 80(4), 409-416.
- Kyriakidis, M., de Winter, J. C., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., . . . Merat, N. (2019). A human factors perspective on automated driving. *Theoretical issues in ergonomics science*, 20(3), 223-249.
- Kyriakidis, M., Happee, R., & De Winter, J. C. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, 32, 127-140.
- Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manufacturing*, 3, 2403-2409.
- Körber, M., Gold, C., Lechner, D., & Bengler, K. (2016). The influence of age on the take-over of vehicle control in highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 39, 19-32.
- Langwieder, K., SPORNER, A., & HELL, W. (1994). STRUKTUR DER UNFAELLE MIT GETOETETEN IM FREISTAAT BAYERN IM JAHR 1991. EIN BEITRAG ZUR ANALYSE DES UNFALLGESCHEHENS.
- Large, D., Banks, V., Burnett, G., Baverstock, S., & Skrypchuk, L. (2017). Exploring the behaviour of distracted drivers during different levels of automation in driving.
- Lehallier, B., Gate, D., Schaum, N., Nanasi, T., Lee, S. E., Yousef, H., Moran Losada, P., Berdnik, D., Keller, A., & Verghese, J. (2019). Undulating changes in human plasma proteome profiles across the lifespan. *Nature medicine*, 25(12), 1843-1850.
- Li, Q., Wang, Z., Wang, W., Zeng, C., Li, G., Yuan, Q., & Cheng, B. (2021). An adaptive time budget adjustment strategy based on a take-over performance model for passive fatigue. *IEEE Transactions on Human-Machine Systems*, 52(5), 1025-1035.

Li, W., Zhang, H., Ni, D., Wang, Z., & Xiao, Y. (2023). The Effect of Driver Fatigue on Takeover Reaction Time under Human-Machine Co-Driving. 2023 7th International Conference on Transportation Information and Safety (ICTIS),

Li, Y., Sharma, A., John, A. P., Alabi, B. N., Chen, S., & Labi, S. (2023). Development of Situational Awareness Enhancing System for Manual Takeover of AV.

Li, S., Blythe, P., Zhang, Y., Edwards, S., Guo, W., Ji, Y., . . . Namdeo, A. (2022). Analysing the effect of gender on the human-machine interaction in level 3 automated vehicles. *Scientific Reports*, 12(1), 11645.

Li, S., Blythe, P., Guo, W., & Namdeo, A. (2018). Investigation of older driver's takeover performance in highly automated vehicles in adverse weather conditions. *IET Intelligent Transport Systems*, 12(9), 1157-1165.

Li, S., Blythe, P., Guo, W., & Namdeo, A. (2019). Investigation of older drivers' requirements of the human-machine interaction in highly automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, 546-563.

Li, S., Blythe, P., Zhang, Y., Edwards, S., Xing, J., Guo, W., Ji, Y., Goodman, P., & Namdeo, A. (2021). Should older people be considered a homogeneous group when interacting with level 3 automated vehicles? *Transportation Research Part F: Traffic Psychology and Behaviour*, 78, 446-465.

Li, S., Blythe, P., Edwards, S., Goodman, P., & Hill, G. (2019). Investigation of the influence of multitasking on drivers' takeover performance in highly automated vehicles. 26th Intelligent Transport Systems World Congress,

Li, S., Blythe, P., Guo, W., & Namdeo, A. (2018). Investigation of older driver's takeover performance in highly automated vehicles in adverse weather conditions. *IET Intelligent Transport Systems*, 12(9), 1157-1165. <https://digital-library.theiet.org/content/journals/10.1049/iet-its.2018.0104>

Lin, Q.-f., Lyu, Y., Zhang, K.-f., & Ma, X.-w. (2021). Effects of non-driving related tasks on readiness to take over control in conditionally automated driving. *Traffic Injury Prevention*, 22(8), 629-633. doi:10.1080/15389588.2021.1969373

Lombardi, D. A., Horrey, W. J., & Courtney, T. K. (2017). Age-related differences in fatal intersection crashes in the United States. *Accident Analysis & Prevention*, 99, 20-29.

- Lonczak, H. S., Neighbors, C., & Donovan, D. M. (2007). Predicting risky and angry driving as a function of gender. *Accident Analysis & Prevention*, 39(3), 536-545.
- Louw, T., Kuo, J., Romano, R., Radhakrishnan, V., Lenné, M. G., & Merat, N. (2019). Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. *Transportation research part F: traffic psychology and behaviour*, 62, 870-882.
- Louw, T. L., Merat, N., & Jamson, A. H. (2015). Engaging with highly automated driving: To be or not to be in the loop? 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design,
- Lyell, D., Magrabi, F., & Coiera, E. (2018). The Effect of Cognitive Load and Task Complexity on Automation Bias in Electronic Prescribing. *Human Factors*, 60(7), 1008-1021. <https://doi.org/10.1177/0018720818781224>
- Ma, S., Zhang, W., Yang, Z., Kang, C., Wu, C., Chai, C., . . . Li, H. (2020). Promote or inhibit: An inverted U-shaped effect of workload on driver takeover performance. *Traffic Injury Prevention*, 21, 1-6. doi:10.1080/15389588.2020.1804060
- Mahdavi, N., Tapak, L., Darvishi, E., Doosti-Irani, A., & Shafiee Motlagh, M. (2024). Unraveling the interplay between mental workload, occupational fatigue, physiological responses and cognitive performance in office workers. *Scientific Reports*, 14(1), 17866.
- Maldonado, T., Orr, J. M., Goen, J. R., & Bernard, J. A. (2020). Age differences in the subcomponents of executive functioning. *The Journals of Gerontology: Series B*, 75(6), e31-e55.
- Manabu, O., Fujioka, T., Hashimoto, N., & Shimizu, H. (2006). The application of RTK-GPS and steer-by-wire technology to the automatic driving of vehicles and an evaluation of driver behavior. *IATSS research*, 30(2), 29-38.
- Marberger, C., Mielenz, H., Naujoks, F., Radlmayr, J., Bengler, K., & Wandtner, B. (2018). Understanding and applying the concept of “driver availability” in automated driving. *Advances in Human Aspects of Transportation: Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation*, July 17– 21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8,
- Marcora, S. M., Staiano, W., & Manning, V. (2009). Mental fatigue impairs physical performance in humans. *Journal of applied physiology*.

Marfoggia, T. T., Heikoop, D. D., Winter, J. C. F. d., & Hagenzieker, M. P. (2020). THE INFLUENCE OF TAKE-OVERS ON DRIVER WORKLOAD: THE ROLE OF PERSONALITY. <https://resolver.tudelft.nl/uuid:d4a53002-3f50-4e4f-af84-62424d8ff15c>

Martens, A. L., Miller, S. S., & Saucier, D. A. (2016). Self-Reports. In V. Weekes-Shackelford, T. K. Shackelford, & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1-5). Springer International Publishing. https://doi.org/10.1007/978-3-319-16999-6_1847-1

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation research part F: traffic psychology and behaviour*, 12(3), 218-224.

McLaughlin, P. M., & Murtha, S. J. (2010). The effects of age and exogenous support on visual search performance. *Exp Aging Res*, 36(3), 325-345. <https://doi.org/10.1080/0361073x.2010.484752>

Medic-Pericevic, S., Mikov, I., Glavaski-Kraljevic, M., Spanovic, M., Bozic, A., Vasovic, V., & Mikov, M. (2020). The effects of aging and driving experience on reaction times of professional drivers. *Work*, 66(2), 405-419.

Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D., Johansson, E., . . . Keinath, A. (2019). The “Out-of-the-Loop” concept in automated driving: proposed definition, measures and implications. *Cognition, Technology & Work*, 21(1), 87-98. doi:10.1007/s10111-018-0525-8

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). Highly Automated Driving, Secondary Task Performance, and Driver State. *Human Factors*, 54(5), 762-771. <https://doi.org/10.1177/0018720812442087>

Miller, D., Sun, A., Johns, M., Ive, H., Sirkin, D., Aich, S., & Ju, W. (2015). Distraction becomes engagement in automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*,

Miller, D., Johns, M., Ive, H. P., Gowda, N., Sirkin, D., Sibi, S., Mok, B., Aich, S., & Ju, W. (2016). Exploring transitional automation with new and old drivers (0148-7191).

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn Psychol*, 41(1), 49-100. <https://doi.org/10.1006/cogp.1999.0734>

- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41(1), 49-100. <https://doi.org/https://doi.org/10.1006/cogp.1999.0734>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*, t. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269.
- Mok, B., Johns, M., Lee, K. J., Miller, D., Sirkin, D., Ive, P., & Ju, W. (2015). Emergency, automation off: Unstructured transition timing for distracted drivers of automated vehicles. 2015 IEEE 18th international conference on intelligent transportation systems,
- Mok, B., Johns, M., Lee, K. J., Miller, D., Sirkin, D., Ive, P., & Ju, W. (2015). Emergency, Automation Off: Unstructured Transition Timing for Distracted Drivers of Automated Vehicles 2015 IEEE 18th International Conference on Intelligent Transportation Systems,
- Mok, B. K., Johns, M., Lee, K. J., Ive, H. P., Miller, D., & Ju, W. (2015, 28 June-1 July 2015). Timing of unstructured transitions of control in automated driving. 2015 IEEE Intelligent Vehicles Symposium (IV),
- Molnar, L. J. (2003). Promising approaches for enhancing elderly mobility.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140. [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)
- Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H., & Olaverri-Monreal, C. (2020). Automated Driving: A Literature Review of the Take over Request in Conditional Automation. *Electronics*, 9(12), 2087. Retrieved from <https://www.mdpi.com/2079-9292/9/12/2087>
- Muslim, H., Itoh, M., Liang, C. K., Antona-Makoshi, J., & Uchida, N. (2021). Effects of gender, age, experience, and practice on driver reaction and acceptance of traffic jam chauffeur systems. *Scientific Reports*, 11(1), 17874.
- Mäntylä, T. (2013). Gender differences in multitasking reflect spatial ability. *Psychological Science*, 24(4), 514-520.
- Mäntylä, T., & Todorov, I. (2013). Questioning anecdotal beliefs and scientific findings: A reply to Strayer, Medeiros-Ward, and Watson (2013). *Psychological Science*, 24(5), 811-812.

- Mäntylä, Karlsson, M. J., & Marklund, M. (2009). Executive Control Functions in Simulated Driving. *Applied Neuropsychology*, 16(1), 11-18. <https://doi.org/10.1080/09084280802644086>
- Mäntylä, T., Carelli, M. G., & Forman, H. (2007). Time monitoring and executive functioning in children and adults. *J Exp Child Psychol*, 96(1), 1-19. <https://doi.org/10.1016/j.jecp.2006.08.003>
- NHTSA., N. H. T. S. A. (2019). Traffic safety facts: Older population. <https://www.nhtsa.gov>
- Naujoks, F., Höfling, S., Purucker, C., & Zeeb, K. (2018). From partial and high automation to manual driving: Relationship between non-driving related tasks, drowsiness and take-over performance. *Accident Analysis & Prevention*, 121, 28-42.
- Nawaz, S., Srivastava, N., Yu, J. H., Baker, R. S., Kennedy, G., & Bailey, J. (2020). Analysis of Task Difficulty Sequences in a Simulation-Based POE Environment. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán, *Artificial Intelligence in Education Cham*.
- Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human factors*, 54(5), 734-746.
- Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors*, 54(5), 734-746.
- Niu, S., Guo, J., Hanson, N. J., Wang, K., Chai, J., & Guo, F. (2024). The effects of mental fatigue on fine motor performance in humans and its neural network connectivity mechanism: A dart throwing study. *Cerebral Cortex*, 34(3), bhae085.
- Ou, Y.-K., Huang, W.-X., & Fang, C.-W. (2021). Effects of different takeover request interfaces on takeover behavior and performance during conditionally automated driving. *Accident Analysis & Prevention*, 162, 106425.
- O'Hanlon, J. F., & Kelley, G. R. (1977). Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In *Vigilance: Theory, operational performance, and physiological correlates* (pp. 87-109). Springer.
- Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *Journal of clinical and experimental neuropsychology*, 33(4), 456-470.
- Pan, H., He, H., Wang, Y., Cheng, Y., & Dai, Z. (2023). The impact of non-driving related tasks on the development of driver sleepiness and takeover performances in prolonged automated driving. *Journal of safety research*, 86, 148-163.

Peng, Q., Wu, Y., Qie, N., & Iwaki, S. (2022). Age-related effects of executive function on takeover performance in automated driving. *Scientific Reports*, 12(1), 5410.

Peng, Q., Wu, Y., Qie, N., & Iwaki, S. (2022). Age-related effects of executive function on takeover performance in automated driving. *Scientific Reports*, 12(1), 5410. <https://doi.org/10.1038/s41598-022-08522-4>

Petermeijer, S., Bazilinsky, P., Bengler, K., & de Winter, J. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied ergonomics*, 62, 204-215. <https://doi.org/10.1016/j.apergo.2017.02.023>

Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Åkerstedt, T., Charles, A., & Bioulac, B. (2005). Fatigue, sleepiness, and performance in simulated versus real driving conditions. *Sleep*, 28(12), 1511-1516.

Pickering, T., Wright, B., Schücker, L., & MacMahon, C. (2024). Active or passive? Investigating different types of cognitive fatigue. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 78(1), 50.

Portron, A., Perrotte, G., Ollier, G., Bougard, C., Bourdin, C., & Vercher, J.-L. (2024). Getting back in the loop: Does autonomous driving duration affect driver's takeover performance? *Heliyon*, 10(3).

Posner, M. I. (1980). Orienting of attention. *Q J Exp Psychol*, 32(1), 3-25. <https://doi.org/10.1080/00335558008248231>

Preusser, D. F., Williams, A. F., Ferguson, S. A., Ulmer, R. G., & Weinstein, H. B. (1998). Fatal crash risk for older drivers at intersections. *Accident Analysis & Prevention*, 30(2), 151-159.

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 5(1), 56. <https://doi.org/10.1186/1479-5868-5-56>

Qazi, A., Hasan, N., Abayomi-Alli, O., Hardaker, G., Scherer, R., Sarker, Y., . . . Maitama, J. Z. (2022). Gender differences in information and communication technology use & skills: a systematic review and meta-analysis. *Education and Information Technologies*, 1-34.

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How Traffic Situations and Non-Driving Related Tasks Affect the Take-Over Quality in Highly Automated Driving.

Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58(1), 2063-2067.
doi:10.1177/1541931214581434

Ranchet, M., Paire-Ficout, L., Uc, E. Y., Bonnard, A., Sornette, D., & Broussolle, E. (2013). Impact of specific executive functions on driving performance in people with Parkinson's disease. *Movement disorders*, 28(14), 1941-1948.

Rauffet, P., Botzer, A., Chauvin, C., Saïd, F., & Tordet, C. (2020). The relationship between level of engagement in a non-driving task and driver response time when taking control of an automated vehicle. *Cognition, Technology & Work*, 22(4), 721-731.

Reimer, B., Donmez, B., Lavallière, M., Mehler, B., Coughlin, J. F., & Teasdale, N. (2013). Impact of age and cognitive demand on lane choice and changing under actual highway conditions. *Accident Analysis & Prevention*, 52, 125-132.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(4), 501.

Riesthuis, P. (2024). Simulation-Based Power Analyses for the Smallest Effect Size of Interest: A Confidence-Interval Approach for Minimum-Effect and Equivalence Testing. *Advances in Methods and Practices in Psychological Science*, 7(2), 25152459241240722.

Robinson, D., Waters, S., Basacik, D., Whitmore, A., & Reed, N. (2015). A pilot study of low workload in train drivers (1910377333).

Roche, F., Somieski, A., & Brandenburg, S. (2019). Behavioral Changes to Repeated Takeovers in Highly Automated Driving: Effects of the Takeover-Request Design and the Nondriving-Related Task Modality. *Human Factors*, 61(5), 839-849. doi:10.1177/0018720818814963

Roche, F., Somieski, A., & Brandenburg, S. (2019). Behavioral changes to repeated takeovers in highly automated driving: effects of the takeover-request design and the nondriving-related task modality. *Human Factors*, 61(5), 839-849.

SAE, & International. (2021, 13/01/2024). SAE Levels of Driving Automation™ Refined for Clarity and International Audience. SAE Blog.

SAE. (2021, 19 March). SAE Levels of Driving Automation™ Refined for Clarity and International Audience. SAE Levels of Driving Automation™ Refined for Clarity and International Audience. <https://www.sae.org/blog/sae-j3016-update>

SAE. (2021). SAE Levels of Driving Automation™ Refined for Clarity and International Audience. Retrieved from <https://www.sae.org/blog/sae-j3016-update>

Sahaï, A., Barré, J., & Bueno, M. (2021). Urgent and non-urgent takeovers during conditional automated driving on public roads: The impact of different training programmes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 81, 130-143.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol Rev*, 103(3), 403-428. <https://doi.org/10.1037/0033-295x.103.3.403>

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology*, 19(4), 532-545. <https://doi.org/10.1037/0894-4105.19.4.532>

Salthouse, T. A. (2010). Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. *Brain and cognition*, 73(1), 51-61.

Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *Journal of experimental psychology: applied*, 19(4), 287.

Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *Journal of experimental psychology: Applied*, 19(4), 287.

Saylik, R., Williams, A. L., Murphy, R. A., & Szameitat, A. J. (2022). Characterising the unity and diversity of executive functions in a within-subject fMRI study. *Scientific Reports*, 12(1), 8182.

Scharfe, M., & Russwinkel, N. (2019). Towards a cognitive model of the takeover in highly automated driving for the improvement of human machine interaction. TC Stewart (Chair), Proceedings of the 17th International Conference on cognitive modelling. Waterloo, Canada: University of Waterloo,

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Collision Analysis & Prevention*, 41(5), 1087-1093.

Schuch, S. (2016). Task Inhibition and Response Inhibition in Older vs. Younger Adults: A Diffusion Model Analysis [Original Research]. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01722>

Seidler, R. D., Bernard, J. A., Burutolu, T. B., Fling, B. W., Gordon, M. T., Gwin, J. T., ... & Lipps, D. B. (2010). Motor control and aging: Links to age-related brain structural, functional, and biochemical effects. *Neuroscience & Biobehavioral Reviews*, 34(5), 721–733.

Shahini, F., Park, J., Welch, K., & Zahabi, M. (2023). Effects of unreliable automation, non-driving related task, and takeover time budget on drivers' takeover performance and workload. *Ergonomics*, 66(2), 182-197.

Shen, X., Wang, C., Zhou, X., Zhou, W., Hornburg, D., Wu, S., & Snyder, M. P. (2024). Nonlinear dynamics of multi-omics profiles during human aging. *Nature Aging*, 1-16.

Shi, E., & Bengler, K. (2022). Non-driving related tasks' effects on takeover and manual driving behavior in a real driving setting: A differentiation approach based on task switching and modality shifting. *Accident Analysis & Prevention*, 178, 106844.

Shuo Li, P. B., Simon Edwards, Paul Goodman, Graeme Hill. (2019). Investigation of the influence of multitasking on drivers' takeover performance in highly automated vehicles

Soares, S., Lobo, A., Ferreira, S., Cunha, L., & Couto, A. (2021). Takeover performance evaluation using driving simulation: a systematic review and meta-analysis. *European transport research review*, 13, 1-18.

Soares, S., Monteiro, T., Lobo, A., Couto, A., Cunha, L., & Ferreira, S. (2020). Analyzing driver drowsiness: From causes to effects. *Sustainability*, 12(5), 1971.

Song, W., Woon, F. L., Doong, A., Persad, C., Tijerina, L., Pandit, P., Cline, C., & Giordani, B. (2017). Fatigue in younger and older drivers: effectiveness of an alertness-maintaining task. *Human factors*, 59(6), 995-1008.

Stanton, N. A., Young, M., & McCaulder, B. (1997). Drive-by-wire: The case of driver workload and reclaiming control with adaptive cruise control. *Safety Science*, 27(2), 149-159. doi:[https://doi.org/10.1016/S0925-7535\(97\)00054-4](https://doi.org/10.1016/S0925-7535(97)00054-4)

Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, 48(10), 1294-1313.

Staub, B., Doignon-Camus, N., Després, O., & Bonnefond, A. (2013). Sustained attention in the elderly: What do we know and what does it tell us about cognitive aging? *Ageing research reviews*, 12(2), 459-468.

- Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing cognitive distraction in the automobile. *Human factors*, 57(8), 1300-1324.
- Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Gender invariance in multitasking: A comment on Mäntylä (2013). *Psychological Science*, 24(5), 809-810.
- Subway princess runner (2023). IvyMobile, L. (Version Version 3.7) [Mobile application software]. Retrieved from <https://apps.apple.com/us/app/subway-princess-runner/id1274410503>
- Thiffault, P., & Bergeron, J. (2003). Fatigue and individual differences in monotonous simulated driving. *Personality and individual differences*, 34(1), 159-176.
- Thompson, P. D., Colebatch, J. G., Brown, P., Rothwell, J. C., Day, B. L., Obeso, J. A., & Marsden, C. D. (1992). Voluntary stimulus-sensitive jerks and jumps mimicking myoclonus or pathological startle syndromes. *Mov Disord*, 7(3), 257-262. <https://doi.org/10.1002/mds.870070312>
- Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *The Quarterly Journal of Experimental Psychology Section A*, 54(2), 321-343. <https://doi.org/10.1080/713755969>
- Tong, Y., Jia, B., Bao, S., Wu, C., & Sethuraman, N. (2024). The difference in physical and mental fatigue development between novice young adult and experienced middle-aged adult drivers during simulated automated driving. *Journal of safety research*, 91, 165-174.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97-136. [https://doi.org/https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/https://doi.org/10.1016/0010-0285(80)90005-5)
- Useche, S. A., Peñaranda-Ortega, M., Gonzalez-Marin, A., & Llamazares, F. J. (2021). Assessing the effect of drivers' gender on their intention to use fully automated vehicles. *Applied Sciences*, 12(1), 103.
- Van der Linden, M., Brédart, S., & Beerten, A. (1994). Age-related differences in updating working memory. *British journal of psychology*, 85(1), 145-152.
- Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience & Biobehavioral Reviews*, 26(7), 849–857.
- Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science*, 20(3), 174-180.

Verhaeghen, P., & Basak, C. (2005). Ageing and switching of the focus of attention in working memory: results from a modified N-back task. *Q J Exp Psychol A*, 58(1), 134-154. <https://doi.org/10.1080/02724980443000241>

Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: a review of meta-analyses. *Neurosci Biobehav Rev*, 26(7), 849-857. [https://doi.org/10.1016/s0149-7634\(02\)00071-4](https://doi.org/10.1016/s0149-7634(02)00071-4)

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: a meta-analysis. *Psychology and aging*, 18(3), 443.

Verissimo, J., Verhaeghen, P., Goldman, N., Weinstein, M., & Ullman, M. T. (2022). Evidence that ageing yields improvements as well as declines across attention and executive functions. *Nature Human Behaviour*, 6(1), 97-110.

Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis & Prevention*, 126, 70-84.

Vogelpohl, T., Kühn, M., Hummel, T., Gehlert, T., & Vollrath, M. (2018). Transitioning to manual driving requires additional time after automation deactivation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 55, 464-482.

WHO. (2021). Decade of healthy ageing: baseline report. World Health Organization.

Wan, J., & Wu, C. (2018). The effects of lead time of take-over request and nondriving tasks on taking-over control of automated vehicles. *IEEE Transactions on Human-Machine Systems*, 48(6), 582-591.

Wandtner, B., Schmidt, G., Schoemig, N., & Kunde, W. (2018). Non-driving related tasks in highly automated driving-effects of task modalities and cognitive workload on take-over performance. Paper presented at the AmE 2018-Automotive meets Electronics; 9th GMM-Symposium.

Weaver, S. M., Roldan, S. M., Gonzalez, T. B., Balk, S. A., & Philips, B. H. (2022). The effects of vehicle automation on driver engagement: the case of adaptive cruise control and mind wandering. *Human factors*, 64(6), 1086-1098.

- Weaver, B. W., & DeLucia, P. R. (2022). A Systematic Review and Meta-Analysis of Takeover Performance During Conditionally Automated Driving. *Human Factors*, 64(7), 1227-1260. doi:10.1177/0018720820976476
- Weiss, H. B., Kaplan, S., & Prato, C. G. (2014). Analysis of factors associated with injury severity in crashes involving young New Zealand drivers. *Accident Analysis & Prevention*, 65, 142-155.
- Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? [Original Research]. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00433>
- Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K., & Connor, J. L. (2011). The link between fatigue and safety. *Accident Analysis & Prevention*, 43(2), 498-515.
- Williamson, A. M., & Feyer, A.-M. (2000). Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication. *Occupational and environmental medicine*, 57(10), 649-655.
- Wong, K. C., & Cheung, W. K. (2012). A study of gender differences in ICT competency. Paper presented at the 2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012).
- Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Factors influencing the latency of simple reaction time. *Frontiers in human neuroscience*, 9, 131.
- World Health Organization. (2015). Global status report on road safety 2015. World Health Organization. <https://www.who.int/publications/i/item/9789241565066>
- Worringer, B., Langner, R., Koch, I., Eickhoff, S. B., & Eickhoff, C. R. (2019). Common and distinct neural correlates of dual-tasking and task-switching: A meta-analytic review and a neuro-cognitive processing model of human multitasking. *Brain Structure and Function*, 224, 1845–1869.
- Wu, H., Wu, C., Lyu, N., & Li, J. (2022). Does a faster takeover necessarily mean it is better? A study on the influence of urgency and takeover-request lead time on takeover performance and safety. *Accident analysis & prevention*, 171, 106647.
- Wu, Y., Kihara, K., Takeda, Y., Sato, T., Akamatsu, M., & Kitazaki, S. (2019). Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving. *Accident Analysis & Prevention*, 124, 202-209.

- Wu, Y., Kihara, K., Hasegawa, K., Takeda, Y., Sato, T., Akamatsu, M., & Kitazaki, S. (2020). Age-related differences in effects of non-driving related tasks on takeover performance in automated driving. *Journal of safety research*, 72, 231-238.
- Wu, H., Wu, C., Lyu, N., & Li, J. (2022). Does a faster takeover necessarily mean it is better? A study on the influence of urgency and takeover-request lead time on takeover performance and safety. *Accident Analysis & Prevention*, 171, 106647.
- Wu, Y., Kihara, K., Takeda, Y., Sato, T., Akamatsu, M., & Kitazaki, S. (2019). Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving [Article]. *Accident Analysis and Prevention*, 124, 202-209. <https://doi.org/10.1016/j.aap.2019.01.013>
- Xu, C., Li, P., Li, Y., Merat, N., Lu, Z., & Guo, X. (2022). Takeover performance and workload under varying automation levels, time budget and road curvature. 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC),
- Yoon, S. H., & Ji, Y. G. (2019). Non-driving-related tasks, workload, and takeover performance in highly automated driving contexts [Article]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, 620-631. <https://doi.org/10.1016/j.trf.2018.11.015>
- Young, K., Regan, M., & Hammer, M. (2007). Driver distraction: A review of the literature. *Distracted driving*, 2007, 379-405.
- Young, M. S. (2021). In Search of the Redline: Perspectives on Mental Workload and the 'Underload Problem'. Paper presented at the Communications in Computer and Information Science.
- Young, M. S., & Stanton, N. A. (2002). Attention and automation: New perspectives on mental underload and performance. *Theoretical issues in ergonomics science*, 3(2), 178-194.
- Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Hum Factors*, 44(3), 365-375. [doi:10.1518/0018720024497709](https://doi.org/10.1518/0018720024497709)
- Yousfi, E., Dogan, E., Bellet, T., Tijus, C., & Guillaume, A. (2018). The effect of time budget and driving assistance on manual takeover performance in a highly automated vehicle. . In *The Human Factors and Ergonomics Society Europe Chapter 2018 Annual Meeting*, 8–10 Oct, Berlin, Germany.

Zeeb, K., Härtel, M., Buchner, A., & Schrauf, M. (2017). Why is steering not the same as braking? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation research part F: traffic psychology and behaviour*, 50, 65-79.

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis & Prevention*, 92, 230-239. doi:<https://doi.org/10.1016/j.aap.2016.04.002>

Zeeb, K., Härtel, M., Buchner, A., & Schrauf, M. (2017). Why is steering not the same as braking? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 50, 65-79. doi:<https://doi.org/10.1016/j.trf.2017.07.008>

Zeeb, K., Härtel, M., Buchner, A., & Schrauf, M. (2017). Why is steering not the same as braking? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 50, 65-79. <https://doi.org/https://doi.org/10.1016/j.trf.2017.07.008>

Zeintl, M., & Kliegel, M. (2010). Proactive and coactive interference in age-related performance in a recognition-based operation span task. *Gerontology*, 56(4), 421-429.

Zelazo, P. D., Craik, F. I. M., & Booth, L. (2008). Executive function across the life span. *Acta Psychologica*, 127(3), 299–307. <https://doi.org/10.1016/j.actpsy.2007.11.002>

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 285-307. doi:<https://doi.org/10.1016/j.trf.2019.04.020>

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 64, 285-307. <https://doi.org/https://doi.org/10.1016/j.trf.2019.04.020>

Zmud, J., Ecola, L., Phleps, P., & Feige, I. (2013). The future of mobility: Scenarios for the United States in 2030.

Zuber, S., Ihle, A., Loaiza, V. M., Schnitzspahn, K. M., Stahl, C., Phillips, L. H., Kaller, C. P., & Kliegel, M. (2019). Explaining age differences in working memory: The role of updating, inhibition, and shifting. *Psychology & neuroscience*, 12(2), 191.

van Loon, R. J., Brouwer, R. F., & Martens, M. H. (2015). Drowsy drivers' under-performance in lateral control: How much is too much? Using an integrated measure of lateral control to quantify safe lateral driving. *Accident Analysis & Prevention*, 84, 134-143.

Özkan, T., & Lajunen, T. (2006). What causes the differences in driving between young men and women? The effects of gender roles and sex on young drivers' driving behaviour and self-assessment of skills. *Transportation Research Part F: Traffic Psychology and Behaviour*, 9(4), 269-277.

Appendix

Appendix A

Table 1. F-statistics for Reaction Time results from ANOVA analysis for experiment one.

<i>Factor</i>	<i>Variable</i>	<i>F-statistics</i>	<i>p-value</i>	<i>Eta-Squared</i>
Task Condition	Reaction Time	65.721 (2.107, 71.627)	< 0.001**	0.659
Age		23.871 (1, 34)	< 0.001**	0.412

*** considered significant for $p < 0.05$; Mauchly's Test of Sphericity < 0.01; Greenhouse-Geisser = 0.701.*

Table 2 Descriptive statistics of the RT results for experiment one.

<i>Factor</i>		<i>Mean RT</i>	<i>SD. Error</i>
Age	Young	1.822	0.099
	Old	2.509	0.099
Task Condition	Updating	1.663	0.062
	Shifting	2.062	0.101
	Inhibition	2.497	0.074
	Combined	2.440	0.083

Table 3. F-statistics for Accuracy results from ANOVA analysis for experiment one.

<i>Factor</i>	<i>Variable</i>	<i>F-statistics</i>	<i>p-value</i>	<i>Eta-Squared</i>
Task Condition	Accuracy	3.451 (3, 102)	0.019**	0.092
Age		6.297 (1, 34)	0.017**	0.156

*** considered significant for $p < 0.05$*

Table 4. Descriptive statistics of the RT results for experiment two.

<i>Factor</i>	<i>Condition</i>	<i>Mean RT</i>	<i>SD. Error</i>
Task Condition	Updating	1.129	0.131
	Shifting	1.614	0.307
	Inhibition	2.200	0.162
	Combined	2.003	0.278

Table 5 F-statistics from ANOVA analysis on RT for experiment two.

<i>Factor</i>	<i>Variable</i>	<i>F-statistics</i>	<i>p-value</i>	<i>Eta-Squared</i>
---------------	-----------------	---------------------	----------------	--------------------

Task Condition	Reaction Time	5.720 (2.342, 44.495)	< 0.001**	0.867
-----------------------	---------------	-----------------------	-----------	-------

**** significant for $p < 0.05$. Mauchly's Test of Sphericity = 0.024. Greenhouse-Geisser = 0.704.**

Table 6. F-statistics for Accuracy results from ANOVA analysis for experiment two.

Factor	Variable	F-statistics	p-value	Eta-Squared
Task Condition	Accuracy	7.768 (2.383, 45.286)	< 0.001**	0.290

**** considered significant for $p < 0.05$. Mauchly's Test of Sphericity = 0.020. Greenhouse-Geisser = 0.781.**

Appendix: A1

Post-Experiment Questionnaire: Takeover Behaviour in Autonomous Vehicles

Simulator Experience

Please read the following questions regarding your experience in the simulator and answer accordingly.

1. Participant ID:

2. Rate your agreement with the statements below. A rating of **1 means you Totally Disagree** with the statement while a rating of **5 means you Totally Agree**.

	1 = Totally Disagree	2 = Somewhat Disagree	3 = Neither Agree nor Disagree	4 = Somewhat Agree	5 = Totally Agree
I think autonomous vehicles are dependable.	☐	☐	☐	☐	☐
I think autonomous vehicles are reliable.	☐	☐	☐	☐	☐
Overall, I can trust autonomous vehicles.	☐	☐	☐	☐	☐
I think using autonomous vehicles can make my driving easier.	☐	☐	☐	☐	☐
I think using autonomous vehicles can improve my driving safety performance.	☐	☐	☐	☐	☐
Overall, I think autonomous vehicles are useful for me.	☐	☐	☐	☐	☐
I think autonomous vehicles are easy to learn.	☐	☐	☐	☐	☐
I think autonomous vehicles are easy to control.	☐	☐	☐	☐	☐
I think autonomous vehicles are easy to understand.	☐	☐	☐	☐	☐
Overall, I think autonomous vehicles are easy to use.	☐	☐	☐	☐	☐
I intend to ride in an autonomous vehicle in the future.	☐	☐	☐	☐	☐

I intend to buy an autonomous vehicle in the future.	☐	☐	☐	☐	☐
I will recommend riding in an autonomous vehicle to friends and family.	☐	☐	☐	☐	☐

3. Rate your agreement with the statements below. A rating of **1 means you Totally Disagree** with the statement while a rating of **5 means you Totally Agree**.

	1 = Totally Disagree	2 = Somewhat Disagree	3 = Neither Agree nor Disagree	4 = Somewhat Agree	5 = Totally Agree
I felt relaxed while riding in the autonomous vehicle.	☐	☐	☐	☐	☐
I am willing to ride in this autonomous vehicle again.	☐	☐	☐	☐	☐

Next

Post-Experiment Questionnaire: Takeover Behaviour in Autonomous Vehicles

Thank you very much for your participation in this study. We appreciate your time and effort.

Prev Done