

Srour Zreik, Rawan (2026) *Take over request design for older drivers in Level 3 automated cars: the interplay of age, non-driving related tasks and culture.* PhD thesis.

https://theses.gla.ac.uk/86302/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

Take Over Request Design for Older Drivers in Level 3 Automated Cars: The Interplay of Age, Non-Driving Related Tasks and Culture

Rawan Srour Zreik

Submitted to the University of Glasgow in fulfilment of the requirements for the Degree of
Doctor of Philosophy

School of Computing Science
College of Science and Engineering
University of Glasgow



June 2026

Abstract

Level 3 automated vehicles allow drivers to disengage from active driving and engage in non-driving-related tasks (NDRTs), while still requiring them to resume control when a Take-Over Request (TOR) is issued. This transfer of control presents a major human–machine interaction challenge, particularly for older drivers, whose cognitive, perceptual, and motor capabilities may affect takeover readiness. Existing research has largely focused on younger populations and has often treated older drivers as a homogeneous group, while giving limited attention to the effects of NDRT characteristics and cultural context on takeover behaviour. This thesis investigates how TORs can be designed to better support older drivers in Level 3 automated vehicles. The first contribution examines how older drivers perceive Level 3 automation and demonstrates that older adults should not be treated as a homogeneous group. Through interviews and participatory investigation, the findings reveal meaningful differences between drivers aged 60–69 and those aged 70+ in their expectations, concerns, and perceived benefits of automation. Building on these findings, the second contribution investigates the effect of TOR source on older drivers' control transitions. A simulator-based study demonstrated that presenting TORs simultaneously on the in-vehicle display and the NDRT device significantly improved takeover reaction times, hazard perception, and TOR usability during lower-demand NDRTs, with the greatest benefits observed in drivers aged 70+. The third contribution presents a cross-cultural and age-sensitive participatory design study exploring expectations and preferences for TOR interfaces among younger and older drivers from two cultural contexts, Israel (low traffic-laws compliance) and the UK (high traffic-laws compliance). The findings reveal differences in communication preferences, modality expectations, and perceptions of automation across both age and cultural groups, highlighting the importance of culturally aware and age-sensitive TOR design. The fourth contribution investigates the role of complete NDRT suspension at TOR onset. A simulator-based study demonstrated that suspension significantly reduced takeover reaction times across both low- and high-demand tasks and across all three age groups (22–32, 60–69, and 70+). Importantly, suspension eliminated age-related disparities observed in the non-suspended condition, with the greatest benefits occurring in drivers aged 70+.

Overall, this thesis contributes a more nuanced understanding of takeover support in Level 3 automated vehicles and argues for adaptive, context-aware TOR systems that better accommodate age-related variability and cognitive demands.

Acknowledgements

Funding This research was supported by the *UKRI Centre for Doctoral Training in Socially Intelligent Artificial Agents*, Grant Number EP/S02266X/1.

This thesis marks the realisation of a long-held dream, one shaped by perseverance, love, and the unwavering support of those who walked this path beside me. To my devoted partner, Nidal, thank you for being my anchor and my refuge. Your belief in me never wavered, even when I faltered. Your strength lifted me, and your quiet sacrifices made this journey possible. To my sweet children, Imad and Jana, your laughter, your light, and your presence reminded me each day of what truly matters. You taught me resilience in its purest form and brought joy even when the weight felt too heavy.

To my parents, your love has been my foundation, your wisdom my compass. You nurtured this dream long before it had a name, and I carry your strength in every step I take.

My deepest thanks to Professor Monika Harvey and Professor Steve Brewster, your guidance, encouragement, and steadfast support shaped not only this thesis but the scholar and person I have become. I am forever grateful for your mentorship.

To the Social AI CDT, thank you for fostering an environment of curiosity, growth, and community. You gave me the tools and the space to grow, not only as a researcher but as a dreamer determined to make a difference.

This journey has been filled with challenges—some visible, others carried quietly—but I met each one with courage, held up by the hands of those who believed in me. This is more than the end of a PhD; it is the beginning of everything I once only dreamed was possible.

Contents

Abstract **1**

Acknowledgements **2**

1 Introduction **11**
1.1 Motivation 11
1.2 Thesis Statement 13
1.2.1 Research Questions 13
1.3 Thesis Walk-through and Research Questions 14

2 Literature Review **16**
2.1 Levels of car automation 16
2.2 An Ageing Driving Population 17
2.2.1 Age-related declines in functional abilities in older drivers 18
2.2.2 Benefits of Driving Automation for Older Drivers 19
2.3 Control Transitions: From Driver-Controlled Driving to Automation-Initiated Control Transitions in Level 3 Vehicles 20
2.3.1 Non-Driving-Related Tasks and Their Impact on Control Transition 22
2.3.2 How Age Impacts the Control Transition Dynamics 23
2.3.3 Culture-Related Differences in Takeover Performance 25
2.4 Designing Take-Over Requests for Older Drivers 26
2.5 Key Research Gaps 27
2.6 Publications & Declaration of Co-Authorship 30
2.6.1 Further Publications by the Thesis Author 31

3 Age Matters: Investigating Older Drivers' Perception of Level 3 Autonomous Cars as a Heterogeneous Age Group (AutomotiveUI '23) **32**
3.1 Introduction 32
3.2 Related Work 33
3.3 Methodology 34
3.3.1 Questionnaire 34
3.3.2 Focus groups 35
3.4 Results 36
3.4.1 **Theme 1: TOR communication** 36

3.4.2 **Theme 2: Incentives to use Level 3 cars.** . . . 37
3.4.3 **Theme 3: Trust and safety.** . . . 37
3.4.4 **Theme 4: Non driving related tasks.** . . . 38
3.5 Discussion and future work . . . 38
3.6 Conclusion . . . 40

4 Where's my TOR?: Evaluating the Effect of Take-Over Request Source on Older Drivers' Control Transitions in Level 3 Cars (MobileHCI '24) 41
4.1 Introduction . . . 41
4.2 Related Work . . . 43
4.2.1 The Challenge of Control Transition For Older Drivers . . . 43
4.2.2 Take-Over Request (TOR) Design For Older Drivers . . . 44
4.2.3 Older Drivers As A Heterogeneous Age Group . . . 45
4.2.4 Summary and Research Gap . . . 45
4.3 Methodology . . . 46
4.3.1 Participants . . . 46
4.3.2 Experimental Design and Variables . . . 47
4.3.3 Materials . . . 47
4.4 Procedure . . . 52
4.5 Results . . . 53
4.5.1 Take-Over Performance . . . 53
4.5.2 NASA-TLX Overall Workload . . . 54
4.5.3 User Acceptance of Advanced Transport Telematics Questionnaire (UAATT) . . . 55
4.5.4 Single Ease Questionnaire . . . 56
4.5.5 Hazard Perception . . . 57
4.5.6 Post Experiment Interviews . . . 59
4.6 Discussion . . . 60
4.6.1 Summary . . . 63
4.7 Limitations and Future Work . . . 64
4.8 Conclusion . . . 65

5 All-inclusive TORs: A Cross-Cultural and Age-Sensitive Design for Take-Over Requests in Level 3 Cars (CHI '25) 66
5.1 Introduction . . . 66
5.2 Related Work . . . 68
5.2.1 The Challenge of Control Transition in Level 3 Cars . . . 68
5.2.2 Take-Over Request (TOR) Design . . . 69
5.2.3 Cross-cultural studies on driving behaviour in Level 3 cars . . . 70
5.2.4 Summary and research gap . . . 71
5.3 Methodology . . . 72
5.3.1 Study design . . . 72
5.3.2 Apparatus . . . 73

5.3.3 Participants . . . 74
5.3.4 Procedure . . . 74
5.4 Results . . . 75
5.4.1 NASA-TLX overall workload . . . 75
5.4.2 Theme Extraction . . . 76
5.4.3 Taxonomy development . . . 82
5.4.4 Synthesis of design concepts . . . 83
5.5 Discussion . . . 85
5.5.1 Designing culturally- and age-sensitive TORs . . . 89
5.6 Limitations and future work . . . 90
5.7 Conclusion . . . 91

6 Examining the Role of NDRT Suspension in Facilitating Safe Control Transitions in Level 3 Cars (conditionally accepted at AutomotiveUI '26) 92
6.1 Introduction . . . 92
6.2 Related Work . . . 94
6.2.1 Non-Driving Related Task Interruption in Level 3 Cars . . . 94
6.2.2 Age-Related Differences in Takeover Performance . . . 95
6.3 Methodology . . . 97
6.3.1 Participants . . . 97
6.3.2 Study Design . . . 97
6.3.3 Apparatus . . . 98
6.4 Procedure . . . 99
6.5 Results . . . 100
6.5.1 Take-Over Performance . . . 100
6.5.2 NASA-TLX Overall Workload . . . 101
6.5.3 Hazard Perception . . . 101
6.5.4 User Acceptance of Advanced Transport Telematics Questionnaire (UAATT) . . . 103
6.5.5 Single Ease Questionnaire . . . 104
6.6 Discussion . . . 104
6.7 Limitations and Future Work . . . 107
6.8 Conclusion and design implications . . . 107

7 Overall Discussion and Conclusion 109
7.1 Overall Conclusion . . . 109
7.2 Overall Contributions . . . 111
7.2.1 Conceptual and Theoretical Contributions . . . 111
7.2.2 Methodological Contributions . . . 112
7.2.3 Empirical Contributions to TOR Design . . . 112
7.2.4 System-Level Contributions: NDRT Suspension . . . 112
7.3 Limitations and Future Directions . . . 113
7.3.1 Ecological Validity and Real-World Deployment . . . 113

7.3.2 Scope of NDRTs and Interaction Modalities 113
7.3.3 Temporal Dynamics and Learning Effects 114
7.3.4 Cultural Generalisability 114
7.3.5 Age as a Behavioural Rather Than Chronological Variable 114
7.3.6 Regulatory and Ethical Considerations 114
7.3.7 Summary 115

Appendices **116**

Appendix A1 **116**

Appendix A2 **119**

Appendix B **119**

List of Figures

2.1 Schematic illustration of the sequence of a take-over process [96] 21

3.1 Participants from the age group 60–65 discussing non-driving-related tasks while shown images of the interior of future AVs (left image: [138]; right image: [74]). 35

3.2 Likert score averages for the dependent variables "perceived safety", "perceived trust" and "perceived usefulness" for each age group. 1=low and 5=high. For non-significant results see Appendix A2. 39

4.1

. 48

4.2 A description of a "Work Zone" hazard. 49

4.3 Experimental setup, with (a) foam sheet with 7 actuators placed on the simulator seat (prior to covering with fabric), (b) a participant wearing the eye-tracker and engaged with an NDRT on the smartphone while Level 3 automation is on, (c) "Automation ON" message on the head-up display and (d) the visual component of the TOR issued on the head-up display (the TOR message says "Attention! Please take control and respond to a vehicle on the left") 50

4.4 NDRTs presented on the smartphone. (a) Mr. Bean video NDRT, (b) Mr. Bean video NDRT during a TOR, (c) reading comprehension NDRT and (d) reading comprehension NDRT during a TOR. 51

4.5 The interaction between Age Group and Condition for each NDRT for the dependent variable Reaction Time. The dots represent mean values. *** - P<0.001, ** - P<0.01, *- P<0.05. 54

4.6 The interaction between Age Group and Condition for each NDRT for the dependent variable Overall NASA-TLX score. The dots represent mean values. *** - P<0.001, ** - P<0.01, *- P<0.05. 55

4.7 UAATT Questionnaire scores of Usefulness and Satisfaction for watching a video NDRT. On the x-axis are the items of the questionnaire. "Usefulness" and "Satisfaction" represent the averaged values. 56

4.8 Single Ease Questionnaire scores for each NDRT measuring ease of detection of the TOR (1= hard to detect, 7= easy to detect). *-P<0.05. 57

4.9 Time To First Fixation on the AOI for each NDRT. The dots represent the mean values. ***-P<0.001, *-P<0.05. 58

5.1 (A) A participant designing the TOR on their personal device and HUD laminate. (B) Design study setup with a design group inside a parked car. (C) Foam sheet with 7 actuators placed on the car seat (prior to covering with fabric) 73
5.2 Overall NASA-TLX workload score for each of the NDRTs. The dots represent mean values. *** - P<0.001, ** - P<0.01, *- P<0.05. 76
5.3 Red light band across the dashboard for the visual component of the take over command. 76
5.4 Risk meter showing hazard criticality, integrated into the HUD takeover command, based on Israeli participants' requirements. 76
5.5 Proposed designs by younger (left) and older (right) drivers from IL and UK respectively. Laminate on windscreen represents Head-Up-Display. Both the type and location of the hazard were conveyed through text. For example: "Take control. Object on the left" (Older driver, UK). Hazard location was indicated on the HUD with a flashing frame, dynamically adjusting as the vehicle approached. 78
5.6 For NDRTs involving handheld devices (Reading (1), Watching a video (2), and Texting (3)), older participants consistently designed a unified interface. The "take over" command was prominently displayed at the centre of the device screen, immediately disabling the NDRT and obscuring the screen with text (e.g., "take over" or "alert"), accompanied by device vibration. 79
5.7 Younger driver's (IL) design showing a meter for hazard criticality. 81
5.8 For NDRTs involving handheld devices (Reading (1), Watching a video (2), and Texting (3)), younger drivers preferred a "take over" command prominently displayed in red either at the top of the device's screen or in the periphery. 82
5.9 Older drivers' taxonomy of TOR features for all the NDRTs including a device (upper tree: Reading, Watching a video and Texting) and Chatting with a passenger (bottom tree). Each branch is labelled with: TOR feature (frequency of appearance). 84
5.10 Younger drivers' taxonomy of TOR features for all the NDRTs including a device (upper tree: Reading, Watching a video and Texting) and Chatting with a passenger (bottom tree). Each branch is labelled with: TOR feature (frequency of appearance). Percentages on the branches indicate the proportion of participants wanting that feature. For example: 92% of the younger participants wanted to include a time budget. 85
5.11 "Disruptive TOR" interfaces. Left: the take over command issued on the NDRT device. Right: the HUD design of the TOR. 87

5.12 "Gradual TOR" interfaces. Left: the take over command issued on the NDRT device. Right: the HUD design of the TOR. 87

6.1 NDRTs presented on the smartphone: (a) reading comprehension NDRT, (b) reading comprehension NDRT suspended by a TOR, (c) a participant from age group 22-32 wearing the eye-tracker and engaging with an NDRT on the smartphone while Level 3 automation is on, (d) the visual component of the TOR issued on the head-up display (the TOR message says "Take control. Pedestrian ahead") 99
6.2 The interaction between Age Group and Condition for each NDRT for the dependent variable Reaction Time. The dots represent estimated mean values. *** - P<0.001, ** - P<0.01, *- P<0.05. 100
6.3 Time to First Fixation and Number of Fixations on AOI for Watching a Video NDRT. 103

7.1 User Acceptance of New Transport Telematics 120

List of Tables

4.1 The four study conditions 47
4.2 Table showing the different dependent variables, their measures and collection method. 48

5.1 Summary of synthesised TOR design concepts based on participant designs, thematic analysis and taxonomies. 86

6.1 Effect of NDRT Suspension on Reaction Time by NDRT Type. 102

7.1 Descriptive statistics and pairwise comparisons between older age groups for perceived trust, safety, and usefulness. 119

Chapter 1

Introduction

1.1 Motivation

As automated vehicles (AVs) progress toward real-world deployment, different levels of automation introduce distinct human–machine interaction challenges. Lower levels of automation (Levels 0–2 [1]) require the driver to remain continuously engaged in the driving task and to supervise the system at all times, while higher levels of automation (Levels 4–5 [1]) aim to remove the human driver from the control loop altogether. Level$\tilde{3}$ [1] automation occupies a critical and transitional position between these paradigms. At this level, the automated system is capable of managing the dynamic driving task under certain conditions, allowing the driver to disengage from active driving and engage in non-driving-related tasks (NDRTs). However, the driver remains a fallback operator and is required to resume control when a system-issued Take-Over Request (TOR) occurs.

This handover of control represents a fundamental challenge for Level$\tilde{3}$ automated vehicles: ensuring that drivers can safely and effectively re-engage with the driving task in unplanned or safety-critical situations. Because responsibility dynamically shifts between the automation and the human driver, at Level 3 automation, the Human–Machine Interface (HMI) is central to the effective delivery of the TOR and the driver's ability to safely resume control. Therefore, TORs must be designed with sensitivity to driver diversity, including age-related differences in cognitive, perceptual, and physical capabilities.

Much of the existing literature on TOR performance in Level 3 automated vehicles has focused on young to middle-aged drivers, often assuming a uniformity in capabilities, preferences, and needs. However, a growing body of research highlights that older drivers—who arguably have the most to gain from vehicle automation due to age-related declines in sensory, cognitive, and motor functions—may face unique challenges during takeover scenarios. These include slower reaction times, reduced attentional flexibility, and higher task-switching costs, which may compromise their ability to safely resume control after NDRT engagement.

This issue is particularly pressing given the global demographic shift toward an ageing population. In the UK, 18% of the population was aged 65 or older in 2016, and this figure is projected to rise to 24.7% by 2046 (ONS, 2017). For many older adults, maintain-

ing access to a car is synonymous with preserving independence, social connectedness, and quality of life [99]. Driving remains their preferred mode of transport, with most car trips undertaken as drivers, and trends indicate that older adults are driving more frequently and over longer distances [58]. However, driving is a complex task that demands the integration of visual, motor, and cognitive functions [76]. Age-related functional impairments can compromise this integration, rendering older adults a vulnerable group with increased susceptibility to driving errors, traffic violations, and collisions. As Level$\tilde{3}$ vehicles begin to shift driving responsibility back and forth between the human and the automated system, it is essential to understand how age-related changes interact with TOR design, particularly in the context of distractions from NDRTs. Prior research suggests that tri-modal TORs (combining visual, auditory, and vibrotactile feedback) can improve detectability and takeover performance by communicating information through multiple sensory channels [85]. Such approaches may be particularly beneficial for older drivers, whose sensory processing and attentional flexibility may decline with age [64]. Designing TOR strategies that effectively support these age-related changes is therefore not only a technical challenge, but also a public health and mobility equity concern.

Existing research on older drivers in automated driving has largely treated them as a single, homogeneous group, implicitly assuming that age-related changes occur uniformly across later life. This thesis challenges that assumption by demonstrating that meaningful differences emerge within older driver populations as people move into later stages of ageing. As drivers approach and pass the age of 70, age-related changes in cognitive flexibility, attentional switching, processing speed, and sensory functioning become more pronounced, directly affecting their ability to disengage from non-driving-related tasks, reorient attention, and respond effectively to Take-Over Requests. In contrast, drivers in their early 60s often retain higher levels of functional capability and driving confidence, leading to different interaction needs and expectations of automation.

By explicitly accounting for intra-generational heterogeneity, alongside the effects of non-driving-related task characteristics and cultural context, this work provides a more nuanced understanding of takeover readiness in Level 3 automated vehicles. The findings demonstrate that treating older drivers as a single group risks obscuring critical differences that have direct implications for TOR design and safety. Taken together, this body of work argues for a shift away from one-size-fits-all solutions toward adaptive, context-aware TOR designs that respond to age-related variability within older populations, as well as broader differences across driver groups. In doing so, the thesis contributes evidence to support the safe and inclusive integration of Level 3 automation into future road ecosystems.

In addition to age-related differences, cultural context may also shape how drivers perceive, interpret, and respond to TORs. Driving behaviours are influenced by culturally shaped attitudes toward risk, rule compliance, trust in automation, and expectations of system communication [148, 56, 48, 136]. Emerging evidence suggests that these cultural differences can influence hazard perception, takeover behaviour, risk appraisal, and interactions with automated systems [144, 56, 27, 148]. However, despite the global

deployment ambitions of automated vehicles, cultural variability remains largely overlooked in TOR and HMI design research [26, 46]. As a result, current understanding of how culturally shaped driving norms influence takeover performance remains limited. Addressing this gap is critical for the development of automated vehicle systems that are not only safe and effective, but also globally deployable, inclusive, and sensitive to diverse driving contexts.

1.2 Thesis Statement

While Level 3 automation allows drivers to disengage from active driving and engage in non-driving-related tasks (NDRTs), it still relies on the human driver to safely resume control when prompted by a Take-Over Request. Current TOR research has largely overlooked the unique needs of older drivers, despite their growing presence on the road and their increased likelihood of benefiting from automated driving. This thesis investigates how TOR design can better support ageing drivers by examining the interaction between age, NDRT engagement, TOR presentation modality and location, and cultural context. Over four experiments, the findings show that older drivers should not be treated as a homogeneous group; that presenting the TOR on the device used for the NDRT improves re-engagement; that the cognitive workload of the NDRT significantly shapes TOR effectiveness; and that driving culture meaningfully influences drivers' design preferences and requirements.

1.2.1 Research Questions

RQ1: How do perceived trust, usefulness, and safety of Level 3 automated driving differ between younger drivers (22–25 years) and older drivers (60–80 years), and across the four older-driver age groups (60–65, 66–70, 71–75, and 76–80 years)?
RQ2: How does the source of TORs (IVIS vs. smartphone+IVIS) affect takeover reaction time, hazard perception, and TOR acceptance among two older driver cohorts under varying NDRT demands?

RQ3: How do cultural (UK vs. IL) and age (young vs. old) differences impact drivers' perception of TORs and their corresponding design requirements in Level 3 automated vehicles?

RQ4: What is the impact of suspending versus continuing NDRTs during a TOR on drivers' takeover performance, hazard perception, and TOR acceptance across three age cohorts and two task demand levels?

1.3 Thesis Walk-through and Research Questions

Building on the research gaps identified in Chapter 2, this section outlines the structure of the thesis and explains how each subsequent chapter addresses the research questions.

Chapter 3: Understanding Age-Related Perceptions of Level 3 Automation

This thesis begins by exploring how drivers of different ages perceive and value Level 3 automation. In this first study, older drivers were treated as a heterogeneous group. Focus groups and questionnaires revealed that motivations to adopt Level 3 cars vary significantly across age subgroups. While the 76–80 age group saw automation as a means to compensate for age-related decline and prolong driving independence, the 60–65 group—despite being classified as "older drivers"—expressed limited interest, citing continued confidence in their driving abilities. In contrast, the 22–25 group expressed enthusiasm based on the appeal of cutting-edge technology. These findings highlighted the heterogeneity of older drivers and the need to go beyond chronological age when designing in-vehicle systems. The study raised important questions about how TOR interfaces should be tailored not only to age-related impairments but also to differences in lifestyle and attitudes. Participants across age groups emphasised the need for TOR systems to interrupt ongoing NDRTs and re-engage the driver. This study helped answer RQ1.

→ This led to the second study, that investigated how TOR source design and NDRT interruption can optimise older driver's performance during control transitions.

Chapter 4: Evaluating the Effect of Take-Over Request Source on Older Drivers' Control Transitions in Level 3 Cars

This study examined whether the presentation source of a Take-Over Request (TOR) influenced older drivers' ability to resume manual control in Level 3 automated vehicles. We conducted a driving simulator experiment with participants aged 60–69 and 70+, who engaged in NDRTs of either low or high cognitive demand. The TOR was presented in two ways: (1) simultaneously on a smartphone—where the NDRT was displayed—and the in-vehicle information system (IVIS), or (2) on the IVIS alone. We measured takeover performance, hazard perception, and user acceptance across conditions. For the lower-demand NDRT, presenting the TOR on both the smartphone and IVIS improved performance, hazard perception, and acceptance, with the strongest gains in the 70+ group. In contrast, no benefits were observed for high-demand NDRTs across presentation method or age group. This work contributes empirical evidence that the effectiveness of TOR presentation formats depends on the cognitive demands of concurrent tasks, offering actionable guidance for designing TOR systems that better support older drivers during control transitions. This study helped answer RQ2.

→ These outcomes motivated the next step—Study 3—to explore how drivers themselves would approach TOR design, especially when accounting for age, varying task types and cross-cultural differences.

Chapter 5: Participatory Design of TOR Interfaces Across Cultures and Age Groups

Informed by the performance limitations observed in Study 2 for the high demand NDRT, the third study took a user-centred design approach. Using participatory design sessions in real vehicles, older and younger drivers from two cultural contexts (UK and Israel) were invited to design TOR interfaces with four types of NDRTs. This approach enabled the collection of rich design rationales, interface expectations, and user-generated features. The analysis yielded a taxonomy of TOR interface elements and revealed shared and divergent priorities: older drivers preferred tri-modal alerts and explicitly supported NDRT suspension as part of TOR delivery, while younger drivers preferred bi-modal TORs and emphasized maintaining control over the NDRT. Israeli participants, influenced by regional risk perceptions, advocated for a "Risk-Aware TOR" that adjusts based on hazard severity. The findings underscored that interface design must be responsive to user context—including age, culture, and task environment—and suggested that user-generated design concepts like NDRT suspension require empirical validation. This study helped answer RQ3.

→ This led to the final study, which directly tested the proposed idea of suspending the NDRT as part of the TOR delivery.

Chapter 6: Empirical Evaluation of NDRT Suspension During TORs

The final study in the thesis experimentally evaluated the user-generated design recommendations from Study 3—namely, suspending the NDRT at the moment of a TOR. Using a driving simulator, the study tested whether task suspension improved takeover performance and hazard perception across two NDRT types (reading vs. video) and three age groups (22–32, 60–69, 70+). Results showed that task suspension significantly improved reaction times across all age groups. However, improvements in hazard perception were limited to the low-demand (video) task. Notably, the 70+ and 22–32 groups exhibited the greatest benefits from task suspension, compared to 60-69 group. The findings reveal that suspension alone is insufficient for high-demand NDRTs, which may require additional strategies (e.g., gradual or guided attention). This study helped answer RQ4.

Chapter 2

Literature Review

This chapter reviews the literature relevant to takeover in Level$\tilde{3}$ automated vehicles, with a particular focus on older drivers. It first introduces levels of vehicle automation and the growing ageing driving population, before examining age-related changes in sensory, cognitive, and psychomotor abilities and the potential benefits of automation for older adults. The chapter then reviews research on control transitions in Level$\tilde{3}$ driving, including the effects of non-driving-related tasks, age, and cultural context on takeover performance. Finally, it examines existing approaches to designing Take-Over Requests (TORs) for older drivers and identifies the key research gaps addressed in this thesis.

2.1 Levels of car automation

Advances in vehicle automation are reshaping the relationship between drivers and vehicles. The Society of Automotive Engineers (SAE) defines six levels of driving automation, ranging from Level 0 (manual driving) to Level 5 (full automation) [1]:

- The most basic level of vehicle automation is referred to as "Level 0 No Automation". At this level, vehicles do not have automated control capabilities for either longitudinal or lateral movement. However, they may feature support systems such as forward collision warning (FCW), blind spot warning (BSW), lane departure warning (LDW), satellite navigation, and night vision enhancement. The human driver is solely responsible for all aspects of driving, including control and environmental monitoring.

- The next stage, Level 1, is termed "Driver Assistance". Vehicles at this level can automate either longitudinal or lateral control, but not both simultaneously. For instance, systems like cruise control (CC), adaptive cruise control (ACC), and lane keeping assistance (LKA) fall into this category. The driver must continue to monitor the environment and maintain full responsibility for driving.

- At Level 2, referred to as partial automation, automation can manage both longitudinal and lateral vehicle control simultaneously. While the driver can remove their

hands from the steering wheel and feet from the pedals, they must remain fully engaged, visually attentive, and ready to take over control at any moment.

- Level 3, defined as "Conditional Automation" by the SAE. While drivers at automation Levels 0 to 2 are required to remain actively engaged in the driving task and continuously monitor the driving environment, Level 3 systems can assume full dynamic driving control, including steering, acceleration, braking, and environmental monitoring. Although the driver must remain present, they are permitted to disengage from the driving task and may engage in non-driving-related activities. However, Level 3 systems are limited in scope and may be unable to handle certain scenarios, such as navigating construction zones or rural roads lacking lane markings or network connectivity. In such cases, the system issues a Take-Over Request (TOR), allowing the driver sufficient lead time to regain control. To maintain safety, the driver is expected to respond and take over vehicle control within the specified time frame.

- SAE describes this as "Level 4 High Automation" in which the automated vehicle can automatically initiate a safe mode even if the drive does not take over control of the vehicle in the time requested.

- The highest level of vehicle automation is referred to as "Level 5 Full Automation". These systems are capable of performing all safety-critical driving functions and monitoring the environment throughout the entire journey, under all conditions. While users may be required to input destination or navigation details, they are not expected to take control of the vehicle at any point during the trip.

This thesis focuses on Level 3 vehicle automation, which presents distinct human–machine interaction (HMI) challenges. At this level, drivers may disengage from the driving task but are expected to resume control at short notice when prompted by the system. This transition from passive monitoring or engagement in non-driving tasks to active driving can be demanding for all drivers—and particularly so for older adults—due to age-related cognitive and physical declines. These factors make the control handover process especially complex and critical to address.

.

2.2 An Ageing Driving Population

A population is generally classified as ageing when individuals aged 65 and over constitute more than 7% of the total population [100]. On a global scale, the share of people aged 65 and above increased from 5.1% in 1950 to 6.9% in 2000, reaching 9.3% by 2020 [51]. This figure is projected to rise further, reaching 15.9% by 2050 and 22.4% by 2100

[51]. In the UK alone, over 4 million drivers were aged 70 or older as of 2013, and projections suggest that by 2050, more than 25% of drivers in the EU will be over 70 years old [141]. Similarly, in the US, drivers aged 65+ are expected to account for 25% of all road fatalities by 2030 [117]. The global shift toward an ageing population underscores the importance of incorporating the needs and capabilities of older drivers in the design and development of autonomous vehicles. Age-related declines in cognitive, visual, and motor functions can affect driving performance and safety, making it essential to ensure that autonomous systems are accessible to and supportive of older users. When designed appropriately, autonomous vehicles have the potential to extend older adults' safe mobility, reduce their reliance on others, and enhance their overall independence and quality of life.

2.2.1 Age-related declines in functional abilities in older drivers

Sensory Function Impairments Sensory impairments in older adults typically involve declines in vision and hearing, with visual impairments being particularly critical for safety due to the heavy reliance on visual information while driving [76, 7]. Ageing can lead to reductions in both static and dynamic visual acuity, as well as contrast sensitivity [7]. A reduced useful field of view (UFOV)—the area from which drivers can extract visual information at a glance—has been linked to slower hazard recognition and elevated crash rates in older drivers [68]. In addition, motion sensitivity deteriorates with age; individuals over 70 exhibit significantly reduced motion sensitivity compared to those under 30 [11]. These impairments can affect driving at night or in low-light environments and may make it difficult to perceive road signs, markings, or traffic signals. Beyond normal ageing, age-related eye diseases like cataracts further reduce visual capability and increase crash risk [76]. Hearing loss due to ageing can also limit the ability to perceive critical auditory cues, such as alarms or honking, further compromising driving safety [76, 7]

Cognitive Function Impairments In addition to sensory decline, age-related cognitive deterioration impacts several domains essential for safe driving, including: attention, perceptual-motor coordination, memory, and decision-making [77, 7, 117]. This aligns with the *frontal ageing hypothesis*, which attributes cognitive decline in older adults to structural and functional deterioration of the frontal lobe [24]. The frontal lobe plays a central role in executive functioning—an umbrella term for higher-order mental processes that govern, coordinate, and regulate other cognitive functions [117]. These executive functions support abilities such as maintaining focus, ignoring irrelevant stimuli, and adjusting responses to meet changing task requirements [106], all of which are critical for safe vehicle operation. For example, Pollastek et al. [123] explored how ageing affects selective attention by conducting studies in both simulated and real-world driving contexts. Their findings showed that older adults demonstrated diminished capacity to focus selectively on relevant stimuli when compared to younger drivers. Divided

attention, which involves managing multiple tasks or processing various streams of information simultaneously [107], also shows significant age-related decline.

.

Fragility and Frailty Lastly, ageing is associated with increased physical fragility and frailty [77]. Fragility increases the likelihood of severe injury from a crash of a given intensity, while frailty reduces the capacity for recovery. Older drivers are at significantly higher risk of serious injury or death in collisions [95]. For instance, drivers aged 70–74 have double the fatality rate of those under 60, and this risk increases fivefold for drivers aged 80 and over [90].

Psychomotor Function Impairments Psychomotor abilities—such as coordination, reaction time, and physical flexibility—are fundamental to vehicle operation and tend to decline with age [29]. Age-related delays in reaction time significantly impact both psychological performance and driving outcomes [92]. Reduced muscle strength may hinder the effective operation of vehicle controls, increasing collision risk [29, 77, 107]. Decreases in muscular strength and joint flexibility can limit the precision and speed of steering, braking, or scanning blind spots—factors that heighten collision risk [29].

2.2.2 Benefits of Driving Automation for Older Drivers

In many countries, driving is central to maintaining independence, mobility, and quality of life in later life [5]. For many older adults, the loss of driving privileges can lead to reduced social participation, increased isolation, and a decline in overall well-being [99]. Automated vehicles (AVs) have the potential to extend safe driving years by compensating for some age-related declines in sensory, cognitive, and psychomotor functions [34]. By taking over routine driving tasks, AVs can reduce the demands of sustained attention, hazard detection, and rapid motor responses, allowing older drivers to focus on higher-level decision making.

Automation can also support older drivers in challenging conditions such as complex traffic environments, high-speed roads, or poor weather, where age-related declines are more pronounced [7, 13]. Features such as adaptive cruise control, lane-keeping assistance, and collision avoidance systems can reduce workload and enhance safety margins. Importantly, Level 3 automation permits engagement in non-driving-related activities during certain phases of travel, potentially increasing productivity and relieving the driver from driving tasks. However, these benefits will only be realised if Level 3 systems are designed with ageing users in mind. This includes providing clear and timely TORs, minimising task-switching demands, and ensuring that interfaces are accessible [32, 91]. When appropriately implemented, automation can help older drivers maintain mobility

and autonomy for longer, while improving safety for all road users.

Taken together, these findings highlight that older drivers are likely to experience unique challenges during control transitions in Level$\tilde{3}$ automated vehicles, particularly when rapid attentional shifts and responses to TORs are required. At the same time, the potential mobility and independence benefits offered by automation make older adults a critical user group for future vehicle systems. Despite this, much existing research has treated older drivers as a homogeneous population and has given limited attention to how age-related functional decline, non-driving-related task engagement, and interface design jointly shape takeover readiness. This thesis addresses these gaps by investigating how TORs can be designed to better support older drivers, while accounting for intra-generational age differences, task demand, and cultural context.

2.3 Control Transitions: From Driver-Controlled Driving to Automation-Initiated Control Transitions in Level 3 Vehicles

In manual driving, intention, attention, and action are closely aligned: drivers maintain continuous situation awareness, choose the timing of manoeuvres, and adapt their actions to traffic and environmental conditions. This self-pacing allows them to prepare physically and mentally, postpone transitions when conditions are unfavourable, and manage cognitive load by limiting secondary tasks during demanding moments. Because the driver controls the initiation point, they can align transitions with their own readiness, mitigating the impact of slower responses or uncertainty. In contrast, Level 3 automation fundamentally changes this dynamic. A defining characteristic of Level 3, as outlined by the SAE [1], is that the vehicle can manage all aspects of the driving task within its operational design domain (ODD), but requires the driver to retake control when the system encounters limitations. When these limits are reached, the system issues a Takeover Request (TOR), initiating a transition from automated to manual control [42, 124, 67]. Unlike Levels 0–2, where drivers remain continuously engaged, Level 3 drivers are permitted to disengage entirely from the driving task and perform non-driving-related tasks (NDRTs) during automation [1, 67, 91].

In an automation-initiated transition, the car sets the timing, not the driver. A TOR can arrive when the driver is cognitively and physically detached from the driving environment. This forces the driver to stop engaging in any ongoing NDRT, reorient their attention to the road, rebuild situational awareness, and execute the necessary control inputs—all within a limited and sometimes short time [33, 44]. The process imposes considerable cognitive demands, particularly when the interrupted NDRT is high in visual, manual, or cognitive load, involving sustained attention, working memory, or visual–spatial processing [116, 168]. The shift from self-initiated to automation-initiated transitions means that takeover success depends not only on the driver's capabilities but

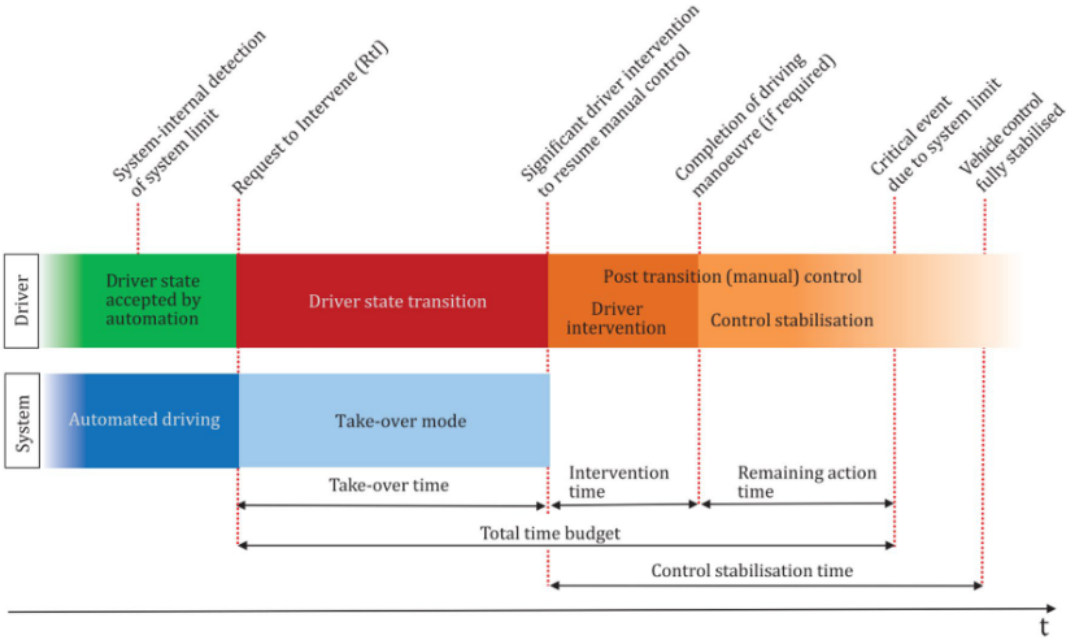


Figure 2.1: Schematic illustration of the sequence of a take-over process [96]

.

also on the design of the TOR—its timing, clarity, and ability to re-engage the driver quickly. These factors are especially critical for older drivers, who may require more time to reorient and act, but they also have a direct impact on takeover safety for drivers of all ages.

Furthermore, system-initiated TORs tend to be more urgent than driver-initiated transitions, as they are triggered by the vehicle detecting environmental or operational limitations (e.g., missing lane markings, construction zones, or sensor failures) [103]. The driver must not only interrupt their ongoing task but also process the context of the takeover situation and execute precise control inputs, such as braking or steering. This sequence of actions requires executive functioning, motor coordination, and perceptual accuracy—abilities that can be compromised by age-related decline or cognitive overload [117, 91, 170]. Control transition during a system-initiated takeover could be concluded in 7 stages as illustrated in Figure 2.1 [96]: **(1)** System detects a system limitation internally, **(2)** Take-Over Request (TOR) is issued (Request to Intervene – RtI), **(3)** Driver begins state transition (shifts attention and prepares to resume control), **(4)** System enters take-over mode (continues automated driving during the transition), **(5)** Driver intervenes (manual input like steering or braking), **(6)** Driver assumes full manual control (post-transition control) and **(7)** Driver stabilises vehicle and completes the necessary manoeuvre.

As such, the control transition in Level 3 vehicles represents a cognitively complex and time-sensitive event. Failure to manage this transition effectively can result in delayed reactions, poor hazard perception, and increased collision risk. This thesis contributes to designing interfaces and strategies to support drivers during takeover events—particularly when engaged in NDRTs—remains a central concern for the safe deployment of conditional automation.

2.3.1 Non-Driving-Related Tasks and Their Impact on Control Transition

At SAE Level 3 automation, drivers are permitted to engage in non-driving-related tasks (NDRTs) while the vehicle handles dynamic driving, but they must remain "fallback ready" to resume manual control when prompted [1]. Understanding the scope and nature of these tasks is essential before evaluating their impact on control transitions. Regulatory bodies have classified NDRTs based on their sensory, motor, and cognitive demands [118], recognizing that these attributes shape their influence on driver performance. Early survey-based research by Schoettle and Sivak [133] revealed that the most common anticipated behavior in a self-driving vehicle was to "watch the road" (41%). Notably, a significant portion of respondents expressed reluctance toward the technology, with 22% stating they would not ride in a fully autonomous vehicle. Among those who mentioned specific activities, the most frequently cited were reading (8%), communicating with friends or family via text or phone (8%), and sleeping (7%). However, such work was limited by reliance on predefined task lists and a lack of real-time behavioral insights. To overcome these limitations, ecologically valid methods have been employed. Shi and Frey [137] compared post-drive self-reports to *in situ* observations in real motorway settings. While participants most commonly reported reading (19.6%), gazing outside (11.8%), and smartphone use (9.8%) as likely NDRTs, observational data revealed different engagement patterns—reading accounted for 30.8% of NDRT duration, followed by smartphone use (18.8%) and gazing outside (16.3%). Unlike self-reports, which capture intent, observational methods provide granular data on duration and frequency, making them more useful for assessing safety relevance in realistic settings.

Theoretical models explain how NDRTs affect takeover performance. Multiple resource theory [59] posits that interference increases when tasks share cognitive and perceptual resources, particularly in visual-manual channels. While this model applies to lower automation levels where driving remains primary, Level 3 reverses this dynamic—NDRTs become the dominant activity, and the takeover represents a task-switching event. Task-

switching theory [81, 137] predicts delays and increased error rates when switching between tasks with dissimilar modalities, especially under time pressure. Thus, the nature of the NDRT and the time allowed for transition thus critically shape takeover success.

Zeeb et al. [169] found that task difficulty (reading vs. proofreading) increased subjective workload, but it was the manual aspect—holding a tablet—that significantly impaired reaction times and lane-keeping performance. Similarly, Lee et al. [89] evaluated nine everyday NDRTs and concluded that cognitive load was the most critical factor in predicting degraded takeover performance. Participants reporting high cognitive effort exhibited more forceful braking and reduced time-to-collision, whereas visual and physical attributes had comparatively less effect. Muller et al. investigated the mental workload of five NDRTs by means of psychophysiological and performance-based parameters as well as the subjective task load. They found that longer reaction times in the NDRT correlate with a higher mental workload and a higher attention allocation. The longest reaction times are found in *Texting* and in *Reading*. *Watching a movie* follows, *Listening* and *Monitoring ride* both have lower reaction times [108].

It is well established that the type of Non-Driving-Related Task significantly influences the success of control transitions in Level 3 automated vehicles. This thesis investigates targeted interventions aimed at mitigating the adverse effects of NDRTs on takeover performance.

2.3.2 How Age Impacts the Control Transition Dynamics

In conventional manual driving, where transitions are almost always initiated by the driver, age-related differences in performance are often moderated by the ability to self-pace. Older drivers are typically aware of their gradual declines in processing speed, attention shifting, and motor coordination, and many adopt conservative driving strategies to compensate [29]. These strategies include maintaining larger headways, avoiding unnecessary lane changes and signalling earlier [29]. Because they control both the timing and the conditions of a manoeuvre, they can often avoid initiating transitions in complex or high-pressure scenarios. This ability to delay action until the environment is more favourable helps mitigate the potential safety impacts of slower cognitive or motor processes.

In contrast, when the transition is initiated by the vehicle in a Level 3 automated driving context, these compensatory strategies become far less effective. The automation determines the timing of the takeover, which may not align with the driver's attentional focus or readiness to act. This externally imposed deadline compresses the time available for reorienting attention, reconstructing situation awareness, and planning a suitable manoeuvre. Older drivers often experience a longer reorientation phase, particularly when a TOR interrupts a demanding non-driving related task such as reading, typing, or using a handheld device [29, 108]. The cognitive costs of task-switching increase with age, meaning that shifting focus back to the driving environment and updating mental models of the situation can take more time [29, 7]. Decision-making under these conditions is

also more demanding. With limited lead time and incomplete situation awareness, older drivers may spend additional moments sampling the environment, double-checking mirrors, and confirming the intentions of surrounding vehicles before committing to an action. While this more deliberate approach helps in reducing decision errors, it can consume valuable seconds when the takeover window is short. As a result, manoeuvres may be initiated later in the available time frame, narrowing safety margins such as time-to-collision. Motor execution presents another challenge. Even when the chosen manoeuvre is correct, the need to compress planning and execution into a short interval can result in more abrupt control inputs—such as higher peak deceleration, sharper steering angles, or greater variability in lateral positioning [92]. These patterns are especially likely in scenarios involving higher speeds, adverse weather, or complex traffic configurations, where the margin for error is already reduced. The combined effect of these factors is a consistent pattern in which older drivers, compared to their younger counterparts, exhibit slower reaction times, longer takeover and indicator times, and less stable vehicle control during automation-initiated transitions. These effects are most pronounced when the TOR arrives during high-load NDRTs [30], in challenging environmental conditions [91], or when the lead time is short [42]. While younger drivers are more likely to respond quickly, they often counteract this advantage by engaging more frequently in distracting tasks [79], which can also impair takeover quality.

Ultimately, the shift from driver-initiated to automation-initiated transitions changes the balance of control and preparation. In manual driving, older drivers can manage transition demands proactively, aligning them with their own readiness and compensatory strategies [29]. In Level 3 automation, however, the lack of control over timing and the need to respond to externally imposed deadlines make age-related slowing and changes in decision-making style more consequential. This dynamic underscores the importance of designing takeover systems that provide sufficient lead time, clear situational cues, and mechanisms to reduce task-switching costs, thereby supporting safe and effective transitions for drivers across all age groups, but particularly for older adults.

Non-Chronological Age Factors and Driving.

2.3.3 Culture-Related Differences in Takeover Performance

Driving behaviours are culturally shaped, influenced by national attitudes toward risk, rule compliance, and enforcement [148, 144, 56, 48, 136]. These behavioural norms not only affect manual driving but also how drivers interact with vehicle automation, particularly in critical scenarios like Takeover Requests (TORs) in Level 3 autonomous vehicles.

A large-scale cross-national study by Özkan et al. using the Driver Behaviour Questionnaire (DBQ) revealed substantial differences in violations and errors across six countries [173]. Countries with higher crash fatality rates—such as Turkey (73 deaths per billion vehicle-kilometres) and Greece (49 deaths)—had significantly higher self-reported violations and aggressive behaviours. In contrast, Finland (7 deaths), the UK (7.5 deaths), and the Netherlands (11 deaths) scored lower on these dimensions, reflecting safer driving cultures [173]. Iran, similarly, presents a high-risk driving environment. Based on a survey of 1,440 drivers, Iranian participants reported more traffic violations, weaker attitudes toward safety, and greater crash involvement compared to Dutch drivers. The fatality rate in Iran is 20.5 per 100,000 population, while in the Netherlands it is just 3.8, ranking among the world's safest [136]. Wang et al. [27] examined how Chinese and German drivers behave in situations involving vulnerable road users, uncovering notable cultural differences in risk response. In two out of three scenarios, Chinese drivers exhibited significantly longer takeover times than their German counterparts, potentially reflecting differing thresholds for perceiving and reacting to risk.

Israel provides a nuanced case: while infrastructure and enforcement are strong, informal norms dominate driver behaviour. Israeli drivers frequently justify rule violations as necessary adaptations to "unfair" conditions, reflecting widespread scepticism toward regulation [135]. Studies comparing Israeli and UK drivers revealed that Israelis predicted traffic hazards more slowly and less accurately—despite equivalent driving experience—indicating differing thresholds of risk perception and attention [148].

In 2018, Israel reported one fatality from road crashes for every 23,440 residents, compared to one fatality per 36,842 residents in the UK [148]. These cross-national patterns have direct implications for human–automation interaction. Hergeth et al. [56] and Strle et al. [144] highlight that driver responses to TORs are not only affected by TOR design and task demands but also by culturally shaped expectations and norms. Hergeth et al. [56] investigated how Take-Over Requests (TORs) influence trust in highly automated driving among Chinese and German drivers. Although Chinese participants reported significantly greater mistrust in automation, trust levels were not found to significantly impact takeover times across the two groups. Similarly, Strle el al. [144] observed that American drivers responded more slowly to TORs than Slovenian drivers, potentially due to greater trust in automation or variations in typical driving behaviour. Despite the global push toward vehicle automation, research on culturally responsive automated vehicle systems remains limited. As noted by Dey et al. [26] and Goodge et al. [46], most interface design efforts have largely overlooked cultural variability. Within this already narrow scope, studies examining control transitions—particularly Take-Over Requests (TORs)— is scarce, especially in non-Western contexts [148]. This gap hinders our understanding of how regional driving norms and culturally shaped risk perceptions influence takeover behaviour, posing a challenge to the global generalisability of current AV design practices.

In sum, national driving cultures shape safety outcomes and driver interactions with AV systems. Designing effective, equitable, and safe Level 3 automation requires nuanced understanding of these cultural patterns—not only to improve TOR performance, but to ensure global usability and trust in future mobility systems. Building on this evidence, this thesis investigates how culturally shaped driving norms influence responses to Take-Over Requests in Level 3 automated vehicles. Specifically, it compares older and younger drivers from the UK and Israel to examine how cultural differences affect preferences for TOR design.

2.4 Designing Take-Over Requests for Older Drivers

Designing effective Take-Over Requests (TORs) is a central challenge in Level 3 automated driving, particularly for older drivers who may experience age-related declines in sensory, cognitive, and psychomotor functioning. Because drivers in Level 3 vehicles can disengage from driving and engage in non-driving-related tasks (NDRTs), TORs must rapidly redirect attention, support situational awareness recovery, and facilitate safe control transitions within a limited time window [140]. Existing research has shown that older drivers generally exhibit slower takeover reaction times, greater workload, and poorer hazard perception during control transitions, particularly when engaged in cognitively or visually demanding NDRTs [117, 91].

Prior work has explored several TOR design dimensions, including modality, urgency representation, timing, message content, and display location. Multimodal TORs

combining visual, auditory, and vibrotactile feedback have consistently demonstrated advantages over uni-modal alerts by improving detectability and reducing takeover time among older drivers [64, 65]. In particular, informative TORs that explain both the need to take control and the reason behind the request have been shown to be preferred by older adults over abstract alerts [92, 128]. These findings suggest that distributing information across multiple sensory channels and providing meaningful contextual information may better support older drivers during control transitions.

The location and source of the TOR have also emerged as important design considerations. Since drivers engaged in NDRTs may direct their attention toward smartphones or infotainment systems, presenting TORs solely through the in-vehicle interface may delay detection and response [103, 157]. Integrating TOR notifications into the NDRT device itself has therefore been proposed as a strategy for reducing attentional switching costs and improving takeover readiness [157]. Beyond TOR presentation, interface accessibility more broadly remains critical for older adults. For instance, older users may find touch-based interfaces less intuitive than younger, more technology-familiar drivers, and often require more button presses to complete non-driving-related activities [50]. Consequently, automated vehicle interfaces should account for the functional limitations and interaction needs of older users [50].

Despite these advances, several limitations remain in the literature. Existing studies have largely focused on younger or mixed-age populations and have often treated older drivers as a homogeneous group, despite evidence that takeover abilities and cognitive flexibility vary substantially within older age cohorts [117]. Furthermore, limited attention has been given to how NDRT cognitive demand, age-related variability, and cultural context jointly shape takeover support needs. Few studies have also incorporated measures such as eye-tracking to capture hazard perception and visual attention allocation during control transitions.

These limitations motivate the work presented in this thesis, which investigates informative tri-modal TORs, the role of TOR presentation source, and adaptive takeover support strategies for older drivers across varying levels of NDRT cognitive demand and cultural contexts.

2.5 Key Research Gaps

Although substantial progress has been made in understanding control transitions and TOR design, important questions remain about how age of the driver, task demands, and cultural factors shape TOR design requirements. The following research gaps highlight limitations in the current literature and provide the foundation for the studies presented in this thesis.

- **Under-representation of older drivers in control transition research**

  While evidence consistently shows that older drivers face unique challenges during Level$\tilde{3}$ control transitions due to age-related declines in attentional switch-

ing, processing speed, sensory functioning, and motor coordination, they remain under-represented in empirical studies of Take-Over Request (TOR) design. Most existing research has focused on younger drivers or mixed-age samples, often assuming that findings generalise across age groups. However, younger and older drivers differ in their cognitive capabilities, technology experience, perceptions of automation, and takeover performance, raising questions about whether the same TOR designs can effectively support all users. From a practical perspective, vehicle manufacturers may favour a one-size-fits-all TOR solution due to its lower implementation complexity and cost. Yet, it remains unclear whether the challenges experienced by younger, younger-old, and older-old drivers are sufficiently similar to justify such an approach. Furthermore, existing studies frequently treat older drivers as a homogeneous population, overlooking meaningful intra-group differences. For example, drivers in their early 60s may demonstrate greater scepticism towards automation and retain relatively high levels of functional capability, whereas drivers aged 70 and above may view automation as a means of maintaining mobility while experiencing more pronounced declines in takeover-related performance. Without understanding both inter-generational and intra-generational differences, TOR designs risk failing to adequately support the diverse needs of future automated vehicle users, potentially compromising both safety and technology acceptance.

- **Limited exploration of the interaction between NDRT characteristics and control transition performance**

  Although it is well-established that Non-Driving-Related Tasks (NDRTs) impair takeover performance—especially those with high visual–manual or cognitive demands—most research has focused on a narrow set of task types. Many studies fail to systematically vary the sensory and cognitive loads imposed by different NDRTs, overlooking how such variations alter attentional reallocation and decision-making during a TOR. Moreover, few studies have investigated whether TORs should be tailored to the characteristics of the NDRT being performed. Existing work has largely overlooked how task demand may influence the effectiveness of different TOR interventions, such as presenting the TOR on the NDRT device itself or suspending the task at takeover onset. As a result, it remains unclear whether the same takeover support strategy is equally effective across tasks with different cognitive demands, particularly for older drivers. This lack of differentiation is problematic given that high-demand NDRTs can exacerbate age-related performance declines, while lower-demand tasks may present opportunities for more effective re-engagement strategies.

- **Overlooking the NDRT device as a potential source of TOR presentation**

  Existing studies have shown that Non-Driving-Related Tasks (NDRTs) impair older drivers' ability to regain control following a Take-Over Request (TOR), yet research on TOR design for this cohort remains limited. Most work has gathered general

TOR design requirements or assessed performance when TORs are issued via the in-vehicle information system (IVIS), implicitly assuming that drivers' visual attention is directed toward the IVIS at the moment of alert. In reality, older drivers engaged in NDRTs may have their visual and attentional resources fully anchored to the NDRT device, slowing their shift back to the driving scene or even preventing TOR detection altogether. This raises a critical and underexplored design question: where should the TOR be presented to effectively capture attention and facilitate a rapid reorientation to the roadway? Early evidence suggests that redundancy -delivering the TOR through both the IVIS and the device hosting the NDRT- may enhance takeover performance. However, the combined effects of TOR location, NDRT cognitive demand, and within-cohort age differences remain largely unexamined, representing a significant gap in designing interfaces that reliably reengage older drivers in Level 3 automation.

- **Under-exploration of cultural influences on takeover behaviour and TOR design**

  Cultural norms strongly influence driving style, hazard perception, compliance with traffic rules, and interpretations of vehicle feedback. Research has demonstrated cultural effects on reaction times, modality preferences, and trust in automation, yet TOR design studies rarely incorporate cultural context. This omission risks developing handover strategies that perform well in one cultural setting but poorly in another, particularly when cultural factors interact with age-related differences. Understanding and embedding cultural considerations into TOR design is critical for ensuring global applicability and acceptance of Level 3 systems.

- **Overlooked potential of NDRT suspension as a system-level intervention**

  Most TOR research has focused on cue timing and modality, with little attention paid to whether the automated system should actively suspend the NDRT at the moment of a TOR. This omission reflects a broader conceptual gap: prior research implicitly assumes that drivers naturally and promptly disengage from NDRTs upon receiving a TOR. Yet, voluntary disengagement is neither instantaneous nor cognitively free — it requires self-regulated attention shifting under time pressure, drawing on executive resources that may already be strained by the ongoing task [**?** 70]. System-initiated suspension — defined as the complete, externally-enforced cessation of the NDRT until the takeover manoeuvre is complete — removes this self-regulatory burden, offering a different interaction mode from self-initiated disengagement. Older drivers have independently proposed such suspension as a desirable safety feature [142], and related work has demonstrated performance gains when TORs are presented via the NDRT device [140]. However, neither approach directly tests task cessation, and no empirical study has evaluated whether the benefits of suspension outweigh its potential switching costs across different task demands and age groups — a significant gap that this thesis directly addresses.

2.6 Publications & Declaration of Co-Authorship

This thesis presents four papers; three out of four are published at top peer-reviewed HCI venues, while one is currently under review at AutomotiveUI.

Chapter 3 constitutes the following paper:

Rawan Srour Zreik, Monika Harvey, and Stephen Anthony Brewster. Age matters: Investigating older drivers' perception of level 3 autonomous cars as a heterogeneous age group. In Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, **AutomotiveUI '23**.

I initiated, designed, and ran the study. I conducted the data analysis, reported and interpreted the data. The work was regularly discussed with my supervisors, Professors Monika Harvey and Stephen Brewster. I wrote the initial draft of the paper and made the subsequent corrections after receiving feedback from my supervisors.

Chapter 4 constitutes the following paper:

Rawan Srour Zreik, Thomas Goodge, Monika Harvey, and Stephen Brew-ster. Where's my tor?: Evaluating the effect of take-over request source on older drivers' control transition in level 3 cars. Proc. ACM Hum.-Comput. Interact., 8(**MHCI**), September 2024.

I initiated, designed, and ran the study. I conducted the data analysis, reported and interpreted the data. My colleague Thomas Goodge assisted in the technical simulator and Psychopy setting. The work was regularly discussed with my supervisors, Professors Monika Harvey and Stephen Brewster. I wrote the initial draft of the paper and made the subsequent corrections after receiving feedback from my supervisors.

Chapter 5 constitutes the following paper:

Rawan Srour Zreik, Monika Harvey, and Stephen Anthony Brewster. All-inclusive tors: Cross-cultural and age-sensitive design for take-over requests in level 3 cars. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, **CHI '25**, New York, NY, USA, 2025. Association for Computing Machinery.

I initiated, designed, and ran the study. I conducted the data analysis, reported and interpreted the data. The work was regularly discussed with my supervisors, Professors Monika Harvey and Stephen Brewster. I wrote the initial draft of the paper and made the subsequent corrections after receiving feedback from my supervisors.

Chapter 6 constitutes the following paper:

Rawan Srour Zreik, Monika Harvey, and Stephen Anthony Brewster. Examining the Role of NDRT Suspension in Facilitating Safe Control Transitions in Level 3 Cars. Under review at **AutomotiveUI 2026**.

I initiated, designed, and ran the study. I conducted the data analysis, reported and interpreted the data. The work was regularly discussed with my supervisors, Professors Monika Harvey and Stephen Brewster. I wrote the initial draft of the paper and made the subsequent corrections after receiving feedback from my supervisors.

2.6.1 Further Publications by the Thesis Author

The author also contributed to related research projects, including facilitating focus groups with Syrian refugees and asylum seekers and consulting on data analysis for a project exploring a social robot to support newcomers in accessing essential services in the UK.

Ashkenazi, S., **Srour-Zreik, R.**, Skantze, G., Stuart-Smith, J., and Foster, M.E. (2026). Participatory Design for Human-Robot Interaction with Syrian Refugees and Asylum Seekers. **In: 2025 International Conference on Social Robotics + AI (ICSR+AI)** (pp. 92-105).

Ashkenazi, S., **Srour-Zreik, R.**, Skantze, G., Stuart-Smith, J., and Foster, M.E. (2026). Multilingual Robots Assisting Newcomers in Public Spaces. In: Workshop on "Robots for Communities" @ **HRI'26**.

Ashkenazi, S., **Srour-Zreik, R.**, Skantze, G., Stuart-Smith, J., and Foster, M.E. (2025). Crossing Borders Together: Co-Designing Robots with Refugees and Asylum Seekers. In: Workshop on "Robots and Autonomous Systems for Refugees, Asylum Seekers, and Migrants (RAS4RASM)" @ **RO-MAN'25**.

Chapter 3

Age Matters: Investigating Older Drivers' Perception of Level 3 Autonomous Cars as a Heterogeneous Age Group (AutomotiveUI '23)

3.1 Introduction

Level 3 automation allows drivers to be physically and psychologically disconnected from driving, allowing them to engage in non-driving related tasks [1]. Yet, they are still required to take over vehicle control when a take-over request (TOR) is issued, should an issue with the automated drive arise. While all drivers are expected to benefit from Level 3 automation, older drivers, in particular, are likely to experience greater advantages from it. Given the increased potential for age-related perceptual, cognitive, and physical declines for older drivers [63], this rapidly growing population could benefit from the driving support offered by Level 3 automated vehicles (AVs). Age-related physical and cognitive deterioration may impair vehicle control. Thus, older drivers may stop driving [69]. However, many studies have found that quitting driving leads to social isolation, loss of independence, and depression [99, 127]. Older drivers may be wary of handing over control to a system they think is less skilled or that they do not understand [22]. Thus, elderly drivers may ignore the mobility benefits of AVs due to their mistrust and poor attitude. Moreover, most research regards older people as a homogeneous group, ignoring variations in the ageing process that may affect driving [93]. Compared to older and more experienced drivers, younger drivers are more likely to engage in non-driving related tasks (NDRTs) while driving, which may increase their crash risk [41]. Whilst this is dangerous in manual vehicles, the introduction of Level 3 automation will enable young drivers to engage in these NDRTs safely, allowing them to be productive and mul-

titask without having to compromise driving safety. Studies show that younger drivers are more willing to use AVs and are more willing to learn about them than older individuals [114]. Possible reasons for this may be that younger drivers have more trust in new vehicle-related technology and perceive AVs to be more comfortable than conventional vehicles [4]. Considering the fast development of Level 3 automated vehicles, far less attention and financial resources have been given to investigating the acceptance, attitudes, and preferences of vehicle automation among older compared to young drivers [110, 22]. Here, we investigate three measures that influence the adoption of AVs, namely perceived trust [78, 21], safety [21] and usefulness [21] and report the initial findings of an online questionnaire (N=250. We recruited 200 older drivers divided equally into four age groups, 50 young drivers aged 22-25 years old) and focus groups. We aimed to answer the following research questions:
RQ1 How do perceived trust, usefulness and safety of Level 3 differ between young and old drivers?
RQ2 What are the differences in perception of Level 3 among older drivers as a heterogeneous age group?
RQ3 What NDRTs are drivers of different age groups willing to engage with while driving a Level 3?

3.2 Related Work

Sufficient initial trust and perceived usefulness are generally regarded as a prerequisite for the mass adoption of AVs [172, 163]. Individual and socio-cultural elements further influence the decision to trust AVs [78] and age is one of the most crucial factors determining trust in automation [171]. Research indicates that younger generations have higher trust in AVs and are more willing to use them [4, 86, 2]. Yet, in a simulator study by Gold *et al.* [43], where participants drove at 120 km/h on a three-lane roadway with Level 3 automation and experienced three take-over scenarios, the driving experience enhanced self-reported trust in automation and older participants rated the vehicle automation more positively than younger drivers [43]. Older people are drawn to technologies that bring evident benefits to their existing lifestyle and they are often resistant if they cannot anticipate potential usefulness [86]. Xiao *et al.* [162] suggest that households under 35 perceive AVs as being more useful than those aged 65 and older [162]. While these studies explore perceived usefulness and trust of older people in AVs, they overlook the variance in lifestyle and mobility needs among older drivers [167]. Therefore, in our study we explore the perceived trust and usefulness of Level 3 cars for older people as a heterogeneous age group.

Safety is a major selling point of autonomous vehicles [34]. However, AVs are often associated with hazards, unpredictability, and loss of control [84]. If drivers perceive AVs to be safe, their willingness to drive them increases [163]. In their survey, Robertson *et. al* [127] suggest that older drivers are less likely to agree that Level 3 cars are safe. Encouragingly, the results of focus groups indicated that older drivers are receptive to

educational strategies and tools to increase their knowledge of this level of automation [127] and potentially increase their perception of safety. Studies have indicated that older subgroups experience more severe sensory, cognitive, and psycho-motor deficits from age 70 forward [77], affecting driving safety. For example, eighty-year-olds may have more noticeable hearing problems [77] which may, for example, impede receiving auditory warnings while driving. Li *et al.* [93] showed that when regaining control of the level 3 automated vehicle, older old people (70 and older) performed worse than younger old people (60–69). They took longer to react to the takeover request and be ready for manual driving, took over vehicle control more slowly, and had a less consistent takeover quality [93]. Therefore, we find it imperative to explore the perceived safety of age-subgroups of older drivers to inform safety-related HMI design of Level 3 cars.

The type of NDRTs drivers engage in differs with age. Wilson *et al.* [156] found that younger participants (45 and younger) were more likely to participate in leisure and social activities, whereas older participants were more likely to want to remain in the driving position and be ready to take back control [156]. Our study provides further insights into the types of NDRTs that older versus young drivers are willing to engage in while in a Level 3 car. The aim is to inform the HMI design of Level 3 cars to facilitate engagement with NDRTs while mitigating the risks involved when a TOR is issued. In the event of a takeover, an older driver engaged in non-driving-related activities may experience a greater mental workload than a younger driver [20], and as a result, may exhibit significantly different driving behaviours than those of the younger driver. Therefore, investigating NDRT engagement in the different age groups is important to the design of TOR.

3.3 Methodology

3.3.1 Questionnaire

We developed and administered an online questionnaire using Qualtrics (`https://www.qualtrics.com/`).

Questionnaire Design. We repurposed questions from previous surveys (see supplementary materials) to uncover the participants perceived usefulness [172], safety [78, 110], and trust [78] in a Level 3 car. The questionnaire collected quantitative data using five-point Likert scale. Based on Weigl et al. and Wilson et al. we explored the social and personal NDRTs that participants would engage in while in a Level 3 car [155, 156].

Participants. We recruited 200 older drivers divided equally into 4 age groups: 60-65, 66-70, 71-75, and 76-80 years old as well as 50 young drivers aged 22-25 years old.

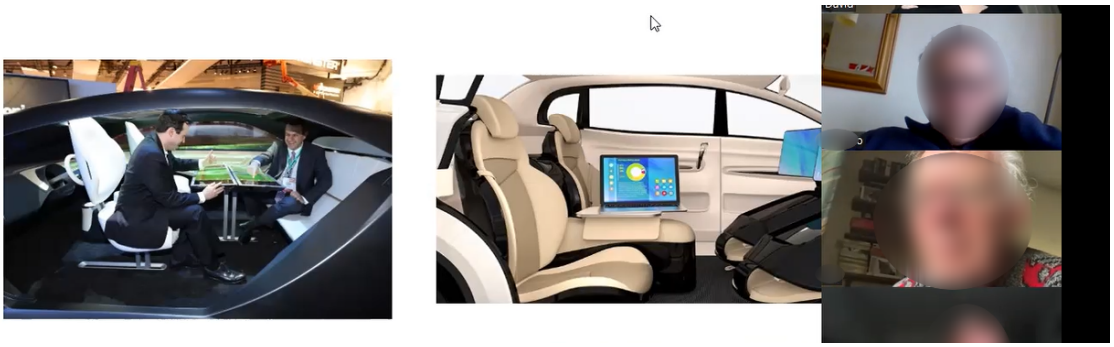


Figure 3.1: Participants from the age group 60–65 discussing non-driving-related tasks while shown images of the interior of future AVs (left image: [138]; right image: [74]).

.

All participants had been active drivers over the past three years and had no previous experience with Level 3 cars. All responses were valid except one from the young drivers' group because the participant failed to answer attention check questions. We received responses from 17 countries while the majority of the respondents were from the UK (n=152).

3.3.2 Focus groups

A series of three focus groups were conducted to expand on key results identified in the questionnaire. All participants in the focus groups had been active drivers over the past three years. The 22-25 group consisted of 3 males and 3 females, the 60-65 group of 4 males and 2 females and the 76-80 group of 3 males and 1 female. Participants were recruited through Prolific platform (`https://www.prolific.co/`). The three focus groups had an identical format, they were held and recorded through Zoom with participants' consent.

Participants were given two driving scenarios: one on a motorway with a TOR after one hour of Level 3 autonomous driving and one near a school with a TOR issued after 10 minutes. They were asked how they felt driving a Level 3 autonomous car in various scenarios, how the TOR should be communicated, what NDRTs they would do, and what information they would like to access when the car is autonomous. See Appendix A1 for the focus groups' questions. Participants were shown four AV interior design images. . The goal was to elicit a discussion on HMI design elements. Participants were compensated with £3.

3.4 Results

Quantitative analysis. We analysed our quantitative data using R statistical tools. We conducted one-way ANOVA with CLM (significance level at 0.05) and *post hoc* comparisons with Bonferroni correction.
Qualitative analysis. Following the guidance of Braun et al. [17] for thematic analysis, we extracted themes from the transcribed focus groups' discussions using NVivo software (`https://lumivero.com/products/nvivo/`). We present and explain the themes supported by quotes from participants and results from the quantitative analysis of the questionnaire.

3.4.1 Theme 1: TOR communication

All participants from the three age groups agreed that the TOR should be communicated through a combination of two or three of the modalities: audio, visual, tactile and it should interrupt any ongoing NDRT. They explained:

> *"the more modalities of interaction or communication the better." (female, 23).*
> *"There needs to be a warning system that overrides whatever you are doing. So, for example if you are on the phone, the call should be interrupted by the warning system." (male, 64).*

Customisability of TOR modalities, displaying a countdown to TOR and automation status (on/off) on the dashboard display were points of consensus between the groups. However, it stood out that the 76-80 group emphasized the importance of a longer lead-time (time from issuing a TOR to handing manual control to the driver) for a TOR for their age group, compensating for the longer reaction time and slower physical movement. Par1 (male,78) stated:

> *"My reaction times are no way near what they were when I was 30. But knowing my abilities have declined I compensate for them. For example, I give myself more time to react. So TOR should be longer for me."*

Comparing scenario 1 with 2, participants across the groups (excluding one participant from the 76-80 group) mentioned that the salience of the TOR should depend on the driving environment. They pointed that the TOR should be more salient in the vicinity of a school because there is a higher chance to have pedestrians and children crossing. Quoting:

> *"More frequent reminders and more prudent ...given the area that you're in." (female, 23).*
> *"[TOR] needs to be louder and firmer because there probably are children around. So it needs to be louder than them. It needs to be more salient and plays more often." (female, 65).*
> *"School area is very busy, I would want frequent warnings" (male,78).*

3.4.2 Theme 2: Incentives to use Level 3 cars.

There were disparities in the incentives of the different groups to using a Level 3. The 22-25 group stated that Level 3 will allow them to use their time more effectively and make journeys more enjoyable. They envision Level 3 cars as being a "mini home office". This group noted that their generation strives to be at the forefront of cutting-edge technologies, which is a significant factor in the adoption of L3 cars. Par2 (female,23) stated:

> *"With our generation we obviously know that we have a significantly longer period ahead and there will be technological advances in different areas. So we have to keep up to date.".*

Our quantitative analysis supports this as the 22-25 group rated the enjoyment while driving a Level 3 significantly higher ($M=3.6$, $SD=1.26$, $p<0.01$) than the 60-65 group ($M=2.68$, $SD=1.42$, $p<0.01$), the 66-70 group ($M=2.62$, $SD=1.39$, $p<0.01$), the 71-75 group ($M=2.71$, $SD=1.23$, $p<0.01$) but not than the 76-80 group ($M=2.88$, $SD=1.25$, $p=0.054$).

The 76-80 group also expressed clear incentives to use Level 3, stating that it will extend their driving lifespan. The decline in vision, delayed reaction times, and the need to compensate for the deterioration of their motor skills are cited as the main reasons for using L3 cars. Par3 (male,79) elaborated: *"My vision is not as good anymore, and I am bothered by glare and lights"*

As opposed to the 22-25 and 76-80 groups, the 60-65 group expressed mainly reluctance towards Level 3 cars. Four out of six participants stated that they enjoy driving and do not have limitations in their capabilities that will prevent them from driving. *"It took me a while to move from manual to automatic. For me it's a pleasure to drive. I would not feel in control in such a car, so I prefer to drive myself."* said Par3 (female,63).

The quantitative analysis revealed similar results. The 60-65 group had the lowest rating of perceived usefulness and differed significantly ($M=2.47$, $SD=0.99$, $p<0.01$) from the 22-25 group which rated usefulness the highest ($M=3.43$, $SD=0.75$) (See Figure 3.2).

3.4.3 Theme 3: Trust and safety.

The themes of trust and safety were interconnected throughout the focus group discussions. All participants indicated that they would be anxious when operating a Level 3 vehicle and would initially have low trust in it. For example, Par3 (male,24) said: *"I think I'd feel anxious because it's not something I've experienced before, so maybe if I'd experienced it a couple of times, I might feel more comfortable."*

The possibility to train on the capabilities of Level 3 and "play" with it using a simulator or a real car was agreed to be a significant step towards establishing trust. Par2 (male,78) commented, *"that would be fantastic to try such technology on a simulator. This allows you to try doing "stupid things". For example, lets see what happens if I do this, in the real road you cannot do trials but in a simulator you can".*

Quantitative analysis revealed that the 22-25 group rated perceived safety significantly higher (M=3.33, SD=0.92, $p<0.05$) than the other age groups except for the 76-80 group (M=2.83, SD=0.98, p=0.082) (see Figure 3.2). The 22-25 group rated that they would feel significantly less anxious (M=2.9, SD=1.07, $p<0.05$) than all age groups except from 76-80 group (M=3.5, SD=1.19, p=0.09) while driving a Level 3 car. There were no significant differences between the groups for perceived trust.

3.4.4 Theme 4: Non driving related tasks.

In agreement with our questionnaire results, the top four NDRTs ranked by participants across the groups of older drivers were: chat to passengers, watch the scenery, listen to music and do nothing. While the 22-25 group rated navigate social media, send emails, listen to music and chat to passengers as the top four NDRTs they would be willing to perform. It could be argued that the different uses of personal technology between the young and older age groups impacted on the non-driving related activities chosen by drivers here. However, one of the selection criteria for both the survey and the focus group was the use of social media platforms; consequently, the older drivers in our study were social media users and it is thus unlikely that the age differences in the expressed NDRT preferences were driven by an overall non-engagement in social media.

Participants across groups agreed that before gaining trust, they are unlikely to engage with highly distracting tasks like reading a book. Par1 (female,64) said: *"I would still be watching the road. I certainly won't read a book, sleep or watch a movie."*

The main concern was that engagement with a highly engaging NDRT may distract the driver from the TOR and compromise safety. Par 3 (male, 79) elaborated: *"I am afraid to miss the take over request and put myself in risk".*

3.5 Discussion and future work

Our study brought interesting insights on the commonalities and disparities in the perception of Level 3 autonomous cars between young drivers and older drivers as a heterogeneous group. We identified that the incentives for using Level 3 cars were polarised. The 22-25 group was enthusiastic about Level 3 cars because they represent cutting-edge technology. The 76-80 group's incentive stems from the acknowledgment of deterioration in vision, cognitive and motor abilities [63]. Therefore, despite concerns about trust, the 76-80 group was keen to use Level 3, as they saw it as an opportunity to extend their driving lifespan. The 60-65 group, however, did not express the same eagerness to use Level 3 automation, stating that it does not answer a pressing need for them as they enjoy driving and perform well at it. **Therefore, a deeper exploration of the lifestyle and driving challenges of this group is needed to make Level 3 cars attractive.**

The questionnaire and focus group found that older drivers were willing to execute NDRTs similar to those they do currently. However, 22-25-year-old drivers were more willing to use social media and send emails. Semi-structured interviews conducted by

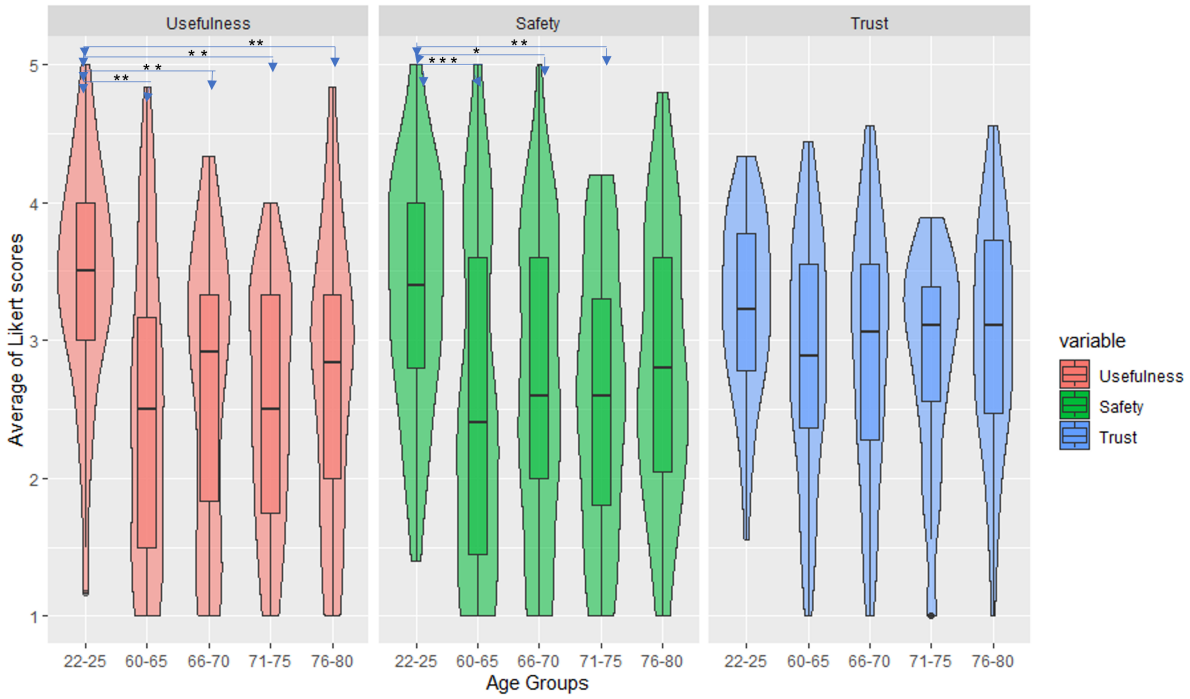


Figure 3.2: Likert score averages for the dependent variables "perceived safety", "perceived trust" and "perceived usefulness" for each age group. 1=low and 5=high. For non-significant results see Appendix A2.

Kaviani et al. [79] revealed that young drivers currently engage with smartphones while driving, which is illegal in many countries and highly dangerous. They perceive smartphone engagement as a natural component of modern life [79]. Level 3, therefore, will enable drivers to engage with smartphones legally and safely. Young and older drivers in the focus group were both reluctant to read or watch a movie. The main problem is that these tasks are excessively distracting, which could hinder manual control after TOR. Design features that facilitate a seamless and safe transition between NDRT and manual control while accommodating age-related physical and cognitive impairments are needed. For example, **how should TOR modalities and intensities be tailored to help 70+ year-old drivers gain situation awareness after task engagement, given their greater cognitive and psychomotor impairments? How should the TOR be designed to interrupt NDRTs for older drivers?**

Lack of initial trust was a commonality between the age groups. As confirmed by Langdon et al. [88] and Gold et al. [43], there was a consensus by participants that first-hand experience with the technology over time, either through a driving simulator or a Level 3 automated car, may improve confidence and trust in the technology [88, 43]. **We advise manufacturers to offer age-specific on-boarding training. When entering the automobile, the system should explain the major features and provide a brief operation manual on the infotainment screen.** Drivers can reduce training as trust grows. Manufacturers could offer Level 3 car demos in their dealerships or PlayStation games for home use. Drivers can test the car's boundaries and make mistakes during training.

3.6 Conclusion

Positioned against existing literature, our study is unique in that it considers older drivers as a heterogeneous group and contrasts their views with younger drivers on Level 3 automation. The results suggest that older drivers should not be regarded as a homogeneous group. The younger old (60-65 group) are more resistant to Level 3 technology, whereas the elder old (76-80 group) view it as a way to maintain their mobility. The communication of the TOR to drivers of various ages should not be uniform, but should take into account the decline in cognitive and physical abilities of those aged 70+ as well as the type of NDRT performed. Level 3's conditional automation introduces a new level of complexity that may cause drivers to become confused and raises safety and trust concerns. If Level 3 HMI design ignores age-related differences in incentives, requirements, cognitive, and physical skills, many potential users may abandon the technology.

Chapter 4

Where's my TOR?: Evaluating the Effect of Take-Over Request Source on Older Drivers' Control Transitions in Level 3 Cars (MobileHCI '24)

4.1 Introduction

According to US official reports, the rate of older drivers on roads (65+) has increased faster than any other age group [131], and by 2030 they will represent a quarter of total driver fatalities – up from around 14.4% in 2002 [95]. In 2013, the UK had more than 4 million drivers aged 70 or more, and by 2050, the EU will have more than 25% of drivers aged over 70 years [131]. Older drivers are more prone to crashes [77, 25], especially at intersections, lane changing and interactions with other road users [102]. Drivers aged 70 and above are more likely to be involved in crashes than those aged 60 to 69 [19, 125].

Driving helps older people stay mobile, which in turn improves their independence and quality of life [141]. However, older adults may cease driving owing to physical and mental deterioration. The development of Level 3 cars is therefore expected to allow them to drive to an older age, thus enhancing their mobility. The Society of Automotive Engineers (SAE) defines six levels of driving automation (Level 0 to Level 5; [1]). Level 0 represents manual driving. Level 1 includes either the longitudinal or the lateral vehicle motion. Level 2 is where the automated system is responsible for both the lateral and longitudinal control of the vehicle. Still, the human driver is the sole driving task supervisor. Level 3 automation enables drivers to be physically and psychologically detached from driving, allowing them to participate in non-driving related tasks (NDRTs) [1]. Yet, when a take-over request (TOR) is issued, they must assume vehicle control. Level 4 is high driving automation, allowing the driver to become a passenger in the car with only minimal intervention and Level 5 represents full automation. The deployment of Levels 4 and 5 is still a distant goal. However, Level 3 vehicles are already being sold in the US

and Germany [53], and the UK Government is in the process of establishing legislation to define responsibility in AV scenarios with and without a human driver capable of taking control [113]. Therefore, research on older drivers' control transition in Level 3 cars is timely.

Assuming manual control after a TOR has been linked closely to the three components of executive functions (Executive Functions (EFs) comprise a set of cognitive processes that control and regulate thought and action. EFs are critical in managing driving tasks [117]): "Inhibition", "Shifting," and "Updating" [106]. In the phase of inhibition, it is vital to suppress the ongoing NDRTs promptly and shift attention to the imminent driving responsibilities [117]. In the phase of shifting, upon receiving a TOR in Level 3 cars, drivers must redirect their cognitive focus from non-driving duties to regaining vehicle control [117]. Finally, in the updating phase, drivers should update their mental model of road conditions to make informed decisions during takeovers [75, 117].

Due to executive function decline with age [24], older drivers may struggle with Level 3 control transitions, compromising safety. Research suggests that Level 3 car take-over requests should therefore consider deteriorating capabilities to meet the needs of older drivers [141, 92, 62, 93, 159]. The existing body of research has focused primarily on gathering requirements for the design of these TORs based on the preferences [92, 141, 128] and performance [93, 64, 158] of older drivers. These studies, however, have only assessed the effectiveness of the TOR when it was delivered solely through the in-vehicle information system (IVIS) (i.e., the visual component of the TOR presented on the HUD or main monitor, the aural component through the vehicle's audio system and the tactile component through the vibration of the seat or seat belt) ignoring the NDRT device (e.g., smartphone or tablet), to which the driver's attention is allocated. This could result in drivers failing or delaying to notice the TOR, compromising safety. We argue that TORs should be incorporated into the NDRT device to assist older drivers in detecting them and switching their attention from the NDRT to the driving task. We conducted a simulator study with older drivers to test our rationale.

In this study, we thus: (1) assessed whether **triggering the TOR on the smartphone and the IVIS ((smartphone+IVIS)-TOR)** can enhance the performance (reduce reaction times and enhance hazard perception) and user acceptance of the TOR for older drivers, as opposed to **triggering the TOR solely by the IVIS (IVIS-TOR)**; (2) investigated differences in performance and user acceptance of TOR between older people from two age groups: 60 to 69 year-olds and over 70 year-old group. Participants were given a smartphone to administer the NDRT and had to respond to multiple TORs during the simulated drive. The two sources ((smartphone+IVIS) vs. IVIS) of the tri-modal (visual, aural and vibro-tactile) TORs were evaluated across two NDRTs varying in their cognitive demand (Higher vs. Lower). Our empirical study makes the following novel contributions:

- We provide evidence that duality of the TOR significantly enhances older drivers' take-over performance for a visual-cognitive NDRT with lower cognitive demand, with greater benefit for the 70+ age group compared to the 60-69 group.

- We demonstrate that duality of the TOR enhances older drivers' hazard perception for a visual-cognitive NDRT with lower cognitive demand, with greater benefit for the 70+ age group compared to the 60-69 group.
- For an NDRT with lower cognitive demand, the duality of TOR significantly enhances usability and facilitates TOR detection for older drivers.
- Lastly, we present recommendations for future research to improve the design of (smartphone+IVIS)-TOR to benefit older drivers while performing NDRTs with higher cognitive demand.

4.2 Related Work

Driver's situational awareness [33] affects the transfer of control from the car to the driver [44, 124, 6]. Prolonged exposure to NDRTs reduces situational awareness below a threshold that is safe for taking manual control, particularly in critical circumstances [154, 117]. A simulator experiment by Zangi et al. [168] involved 37 drivers driving under Level 2 automation with and without a visual-manual NDRT (reading text and answering multiple choice questions on the in-vehicle display). Eye-tracking data showed that drivers were less aware of road hazards when they engaged with the NDRT, and NASA-TLX scores showed that NDRT engagement induced a higher mental workload. Another simulator study by Yoon et al. [165] compared reaction time and time to first fixation on the road from TOR onset across various NDRTs. They found that NDRT type affected takeover reaction time, with smartphone use delaying take-over reaction times more than watching a video. This prompted researchers to focus on TOR design to bring drivers safely back into the driving loop. The current consensus suggests that multi-modal alerts [165] result in the quickest braking response time [122, 67], hands-on steering response time [165], and automation disengagement time (as measured by button presses).

4.2.1 The Challenge of Control Transition For Older Drivers

The challenge of control transition in Level 3 cars is accentuated for older drivers. Age-related cognitive and sensory declines may impair older drivers' ability to respond quickly and effectively to driving alerts after engaging with an NDRT [117], making TOR research critical for this cohort of drivers.

Existing research on older drivers found that they made fewer transitions between hazard areas of interest and had shorter mirror fixation duration when engaged with an NDRT [116]. They also had longer reaction and takeover times and left smaller distances from the vehicles ahead compared to their young counterparts [91]. Wu et al. [158] discovered that takeover performance declined for older drivers (mean age= 64.3) especially when switching between multiple NDRTs. Additionally, according to Ebnali et al. [30], cognitively demanding NDRTs are more dangerous for older drivers due to cognitive

decline in those 65 and older. Despite these challenges, older drivers expressed willingness to perform tasks of varying cognitive demand while in Level 3 autonomous cars, e.g. listening to music, watching videos, chatting to passengers, and reading [92, 141]. Acknowledging the deterioration in their cognitive and physical abilities, older drivers advocated that ongoing NDRTs should be interrupted by the TOR on the NDRT device to facilitate control transition [92, 141]. The aforementioned studies [116, 30, 91, 158] confirm that the adverse effects of NDRT engagement are amplified for older drivers. Yet, older drivers are underrepresented in research on user interface design in control transitions. In their literature review, Happee et al. [82] identified 25 studies that explicitly studied the effectiveness of user interfaces in control transitions in Level 3 cars. Yet, none of these studies included older drivers. This makes it imperative to develop TORs customized for older drivers, carefully considering the demands of NDRTs, to assist them in smoothly and safely transitioning from non-driving activities to driving.

4.2.2 Take-Over Request (TOR) Design For Older Drivers

Numerous studies have investigated the TOR design requirements of older drivers. Huang et al. [64] indicated that older drivers find tri-modal (i.e., visual, aural and vibrotactile) TORs the easiest to detect with faster reaction times compared to uni-modal or bi-modal ones. For the auditory component of the TOR, older drivers prefer speech-based over non-speech-based stimuli due to their ability to give more information [128]. According to Li et al. [92], older drivers prefer a takeover request that informs them about their need to take control of the vehicle and explains the reason for requesting it. Additionally, any alert type that included a tactile component (i.e., auditory-tactile, visual-tactile, and visual-auditory-tactile) was associated with shorter brake response times for older drivers compared to the corresponding alert without the tactile component [65]. These studies, however, conclude with several gaps that we addressed in this paper. The TOR design requirements were not evaluated and refined to coordinate with different NDRTs which are known to affect control transitions in Level 3 [117, 158, 91]. This is particularly critical for subgroups of older adults, whose cognitive abilities, and reaction time performance, may vary [117]. Additionally, Li et al. [92] stress the necessity of evaluating the effectiveness of the TOR presented at different locations, particularly on the device used for performing the NDRT.

In a simulator study by Melcher et al. [103], 44 drivers were distracted by a quiz game on a mobile phone and had to take-over control following the TOR. Two locations for the TOR presentation (mobile phone and vehicle HMI, or vehicle HMI alone) were evaluated. Results showed that including the visual component of the TOR on the driver's mobile phone, increased acceptance of the TOR and perceived safety without affecting response times. Similarly, Wintersberger et al. [157] introduced "Attentive User Interfaces" (AUIs) to ensure safe and efficient use of mobile devices during automated driving. Their study found that a TOR with notifications on the NDRT device led to reduced stress, increased trustworthiness, and a tendency for enhanced takeover. Notably,

however, these studies did not include an older cohort and did not explore the varying levels of cognitive demand associated with NDRTs. Additionally, they did not collect eye-tracking metrics, which are essential for quantifying hazard perception and visual attention allocation. Furthermore, they used abstract rather than informative TORs (i.e., providing the reason for the take-over request). In contrast to informative TORs, abstract TORs can cover a smaller body of semantics and were less preferred by older drivers [92]. Moreover, these studies included only visual [103] or visual-auditory [157] TORs on the smartphone, ignoring vibrotactile feedback.

These shortcomings in the literature motivated us to compare, for the first time, the effectiveness of two sources of an informative tri-modal TOR on older drivers while they performed NDRTs of varying cognitive demand. We proposed and evaluated a solution to mitigate the challenge faced by older drivers in control transitions, taking into consideration the driver's age group and the cognitive demands of the NDRT.

4.2.3 Older Drivers As A Heterogeneous Age Group

The age threshold of 60 years is widely recognised as the standard for categorising individuals as older adults [93, 7]. This cutoff marks the onset of deterioration in hazard perception ability [60, 37, 7] and has also been extensively integrated into previous studies on advanced driver-assistance systems (ADAS) and vehicle automation systems for older adults [87, 20]. However, the elderly population is not a homogeneous group [5, 141, 25], and disregarding the variability inherent in the ageing process would contribute to disparities among aged individuals belonging to distinct subgroups [5, 93]. For example, in regaining control of a Level 3 automated vehicle, Li et al. [93] found that drivers aged 70 and older performed worse than 60–69 year-olds: they were slower to respond to and complete takeover requests. Age also increases road accident casualty rates and fatal injury risk for people 70 and older [25]. The effects of age 70 on driving performance have been noted. In the UK, driving licenses expire at the age of 70, and individuals must renew their licenses every three years to continue driving [28].

These studies have cemented the need for a more nuanced design of TORs tailored for the support of specific sub-groups of older people. Failing to do so, Level 3 cars may be designed following a 'one size fits all' approach, ignoring the capabilities, needs, and requirements of specific subgroups of older people.

4.2.4 Summary and Research Gap

Thus far, it has been established that NDRTs affect older drivers' ability to successfully retake control after a TOR. Yet, older drivers are underrepresented in studies exploring user interface design in control transitions [82]. To ensure a safer deployment of Level 3 vehicles for this growing cohort, it is therefore imperative to optimise TOR design to mitigate the challenges older drivers face during control transitions. Existing research has focused on gathering TOR-design requirements [92, 141, 128, 63] and evaluating older drivers' driving performance after a TOR is issued by the IVIS [93, 62]. These studies

overlooked the possibility that the driver's visual and attentional resources may be devoted to the NDRT device, causing them to transfer their focus slowly or even fail to notice the TOR entirely. Where should the TOR be presented to older drivers to successfully shift their attention from the NDRT back to the driving scene? To fill these gaps, our research goes beyond merely including older drivers; we explore how the duality of the TOR presentation impacts reaction time, hazard perception and TOR acceptance across two distinct groups of older drivers. This approach is crucial given the findings from Li et al. [93] that drivers aged 70 and older perform worse than those aged 60–69 in regaining control of Level 3 automated vehicles.

In our study, we triggered the TOR through either the IVIS alone or on both the IVIS and a smartphone where the NDRT was located. Older drivers were divided into two groups: 60-69 and 70+ years old and engaged with two NDRTs varying in their cognitive demand (Higher vs. Lower).

In this paper, we address the following research questions:

- **RQ1** How does a (smartphone+IVIS)-TOR affect takeover reaction time compared to IVIS-TOR across two cohorts of older drivers, considering two NDRTs with varying levels of cognitive demand?
- **RQ2** How does (smartphone+IVIS)-TOR affect hazard perception compared to IVIS-TOR across two cohorts of older drivers, considering two NDRTs with varying levels of cognitive demand?
- **RQ3** Does (smartphone+IVIS)-TOR enhance TOR acceptance and detection compared to IVIS-TOR across two cohorts of older drivers, considering two NDRTs with varying levels of cognitive demand?

4.3 Methodology

4.3.1 Participants

We recruited 22 healthy older drivers from the UK divided equally into two age groups: 60-69 (Mean age= 64 years, SD= 2.3 years; 6 females and 5 males) and 70+ years old (Mean age= 75 years, SD=3.3 years; 6 females and 5 males). Recruitment was done via internal university e-mail lists and from church communities. All participants were retired and had been active drivers for at least the past 3 years. They reported normal to corrected-to-normal vision and hearing, could see clearly up to a distance of 1 meter (the distance from the simulator screen) and had no compromised sense of touch. All participants declared they knew about the concept of automated cars but had never experienced them. In the 60-69 age group: 1 participant completed education up to high school, 2 had diplomas, 5 held a bachelor's and 3 obtained higher degrees. In the 70+ age group: 4 participants completed education up to high school, 5 held a bachelor's

degree and 2 obtained higher degrees.

One of the participants (in the 60-69 age group) took part in only one session of the experiment, therefore their data were discarded and another participant recruited instead. At the end of the second session, we administered the Montreal Cognitive Assessment (MoCA) [109] to ensure that our participants did not experience signs of cognitive impairment which could interfere with their performance. Early indicators of moderate cognitive impairment, which is a transitional phase between typical ageing and dementia, are identified by the MoCA [109]. A score of 26 or higher signifies intact cognition. All of our participants scored 26 or more (M=27.69, SD=1.3, Min=26, Max=30).

Our Institution's Research Ethics Board granted ethics approval before the research was conducted.

4.3.2 Experimental Design and Variables

The experiment used a 2 (cognitive demand of the NDRT) x 2 (sources of TOR) x 2 (age groups) full factorial design. Cognitive demand levels and sources of TOR were within-subject independent variables and age was a between-subjects independent variable.

Levels of cognitive demand (Higher vs. Lower): participants read passages and answered multiple choice questions (higher cognitive demand NDRT) or watched a video (lower cognitive demand NDRT).

Sources of TOR: tri-modal TORs were triggered either solely by the In-Vehicle Information System (IVIS) or in combination with triggering on the smartphone ((smartphone+IVIS)-TOR).

The combination of NDRT cognitive demand and TOR source resulted in four conditions. Each participant performed all four conditions in a counterbalanced order. See Table 4.1.
Dependent variables are presented in Table 4.2.

		Cognitive demand of NDRT	
		Higher (reading comprehension)	**Lower** (watching a video)
TOR source	**IVIS-TOR**	Condition 1: Passages set 1	Condition 2: Tom and Jerry
	(smartphone+IVIS)-TOR	Condition 3: Passages set 2	Condition 4: Mr. Bean

Table 4.1: The four study conditions

4.3.3 Materials

The participants were seated in a driving simulator (

. The NDRTs were displayed on a Google Pixel

Dependent Variable	Measure	Data collection method
Hazard perception	Time in seconds from onset of TOR until the first fixation (>100 ms) on the AOI	Eye tracker
	Hazard Detection. At least one fixation (>100 ms) that a participant made at the AOI	Eye tracker
	The total number of fixations (>100 ms) that a participant made at the AOI	Eye tracker
Take-over performance	Reaction time. Time in seconds from onset of TOR until the button press (manual takeover)	Button on steering wheel
Subjective ratings of mental workload	6-item, 0-10 scale	NASA-TLX Questionnaire
Subjective ratings of user acceptance of TOR	9-item, (-2) to (2) scale	User Acceptance of New Transport Telematics Questionnaire
Ease of detection of the TOR	Single item, 1 to 7 scale	Single Ease Questionnaire

Table 4.2: Table showing the different dependent variables, their measures and collection method.



Figure 4.1:

7 smartphone provided by the experimenter. The experiment was conducted using PsychoPy v2021.2.3 [115] and presented on an Asus VX279 27-inch monitor positioned at a distance of 1 metre from the participant. See Figure 4.3a.

Driving Scenarios

Multiple drives around the city were recorded from the driver's perspective by affixing a GoPro Hero 360 Max camera to the windscreen of a Citroen C3 vehicle. Video recordings were obtained at different times during the day, from March to May, in and around the vicinity of Glasgow city. This footage was then reviewed and edited into 4 different driving scenarios that lasted between 11 to 13 minutes each. The four scenarios included various hazards that occurred naturally on the road during filming and were not prearranged (the hazards were: road work zone, car pulling out from side road, pedestrians crossing, cyclist approaching and car protruding into the driver's lane). For each driving

scenario we picked 6 of these naturally occurring hazards to issue a TOR. The definition for a hazard was adopted from the UK Government's Hazard Perception test [28] and from [168, 149, 47]. An Area of Interest (AOI) was determined for each hazard. AOI refers to a specific region or zone within the driver's field of view that is particularly relevant for detecting and responding to potential hazards [168]. See Figure 4.2 for an example.

Parked cars obscuring a work zone



Description: The participant's vehicle drives on a one-way lane. Parked cars on the right obscure a work zone.

Hazard: Roadwork equipment. If the participant does not notice the road work signs, the road work ahead may pose a threat.

AOI: "Work zone" ahead signs.

Figure 4.2: A description of a "Work Zone" hazard.

Non-driving Related Tasks

While the vehicle was in "automated" driving mode, participants were required to complete one of two visual-cognitive NDRTs varying in their cognitive demand (depending on the study condition). A mobile application was developed utilising the Flutter open source framework to facilitate engagement with NDRTs and issuing of the TOR. In their research, Muller et al. [108] indicated that reading was significantly more mentally demanding than watching a movie. Therefore, two versions of the lower cognitive demand NDRT were created, one using a video of a Tom and Jerry cartoon [1] and one with a video from a Mr. Bean movie [2]. For the higher cognitive demand NDRT, we chose a collection of reading comprehension passages and multiple-choice questions given to 12 year-old pupils in the USA [3].

Two sets of passages and questions with comparable difficulty were chosen for the 2 versions of this NDRT type.

We evaluated the perceived workload of our NDRTs using the NASA-TLX questionnaire. Twenty older drivers (divided equally into the two age groups and screened with the same criteria as in our study) completed NASA-TLX after engaging with each of the two NDRTs. A linear mixed model was fitted to predict the main effects of NDRT and Age Group on the NASA-TLX scores with a random intercept for each participant. We found a significant main effect of NDRT on the overall NASA-TLX score ($F(1,18)=30$, $p<0.0001$)). Participants rated the Reading NDRTs significantly higher (EM=21.1, SE=2.02) than the Video NDRTs (EM=10.9, SE=2.02). The interaction between NDRT and Age Group was not significant ($F(1,18)=1.2$, $p=0.28$).

[1] Tom and Jerry video `https://www.youtube.com/watch?v=ATW6RvBbn9o`

[2] Mr. Bean video `https://www.youtube.com/watch?v=G9nqB8BwGHU`

[3] Reading comprehension passages `https://englishforeveryone.org/Topics/Reading-Comprehension.html`



Figure 4.3: Experimental setup, with (a) foam sheet with 7 actuators placed on the simulator seat (prior to covering with fabric), (b) a participant wearing the eye-tracker and engaged with an NDRT on the smartphone while Level 3 automation is on, (c) "Automation ON" message on the head-up display and (d) the visual component of the TOR issued on the head-up display (the TOR message says "Attention! Please take control and respond to a vehicle on the left")

Take-Over Request Design

Prior research suggests that take-over requests should be multimodal [63, 122, 67]. Thus, we presented TORs visually, aurally and using vibrotactile feedback. For each of the six hazards, a TOR time budget was defined during which the participant should identify the hazard and regain manual control to avoid a potential crash. The start of the TOR time budget was called the TOR onset and the end as the TOR end. In the literature, the duration of the TOR budget ranges between 6 seconds [149] to 20 seconds [93], for our study the time budget was 10 seconds, following [103].

TOR triggered by the In-Vehicle Information System (IVIS-TOR). The visual and auditory components of the TOR issued by the IVIS were designed following the recommendations of Li et al. [92] for age-friendly TORs and of Hollander et al. [57]. For the auditory component, Logitech speakers were placed on the floor either side of the participant facing towards them. The visual component consisted of a translucent written message overlaid onto the video as a heads-up display providing the reason for the manual driving takeover request. For example, "Attention! Please take over vehicle control and re-

spond to a parked car ahead" [92, 93]. Following a pilot study (N = 4), the message was shortened to "Attention! Please take control and respond to [the hazard]" as the original message was deemed to be too lengthy. As recommended by Hollander et al. [57], we added a dynamic countdown bar and a digital countdown indicating time left until the end of the TOR as a numerical value in seconds (see Figure 4.3d). For the auditory component, a computer-generated (through text-to speech generating software) female voice read the message as done by Li et al. [92]. For the vibrotactile component, 7 actuators from the Actronika HSD MK-I prototyping KIT [4] were embedded in a foam sheet and placed on the seat of the simulator, covered with a fabric (see Figure 4.3a). Vibration was set at a frequency of 250 Hz for 1 second and all actuators vibrated simultaneously. The duration and source of vibrotactile feedback vary in the literature. Yoon et al. [165] used 2-second feedback through the simulator seat, while Huang et al. [64] administered 1-second feedback through the seat belt. We opted for 1-second vibrotactile feedback on the seat-pan to strike a balance between effectively alerting the driver and minimizing the annoyance often associated with multi-modal TORs [121, 111].

TOR triggered by the smartphone. The tri-modal take-over request that was issued by the smartphone included the same visual and auditory message as in the IVIS. For the vibrotactile feedback the phone vibrated for 1 second during the TOR, using the internal actuator. The visual message appeared on the top of the screen in bold red text, with the dynamic bar and digital countdown to the right. See Figure 6.1 (b) and (d). The visual message did not obscure the task and appeared at the top of the screen to give participants the option to manage the interruptibility of the task [157]. Both tasks were presented while the phone was in portrait orientation to ensure uniformity of the visual message display.

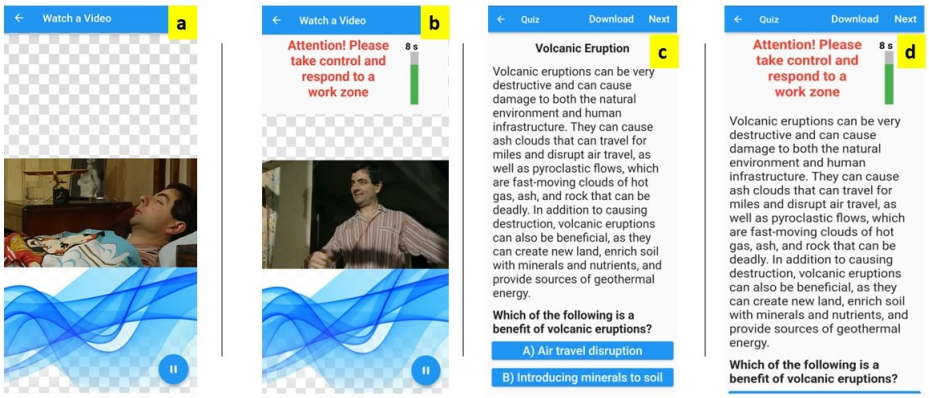


Figure 4.4: NDRTs presented on the smartphone. (a) Mr. Bean video NDRT, (b) Mr. Bean video NDRT during a TOR, (c) reading comprehension NDRT and (d) reading comprehension NDRT during a TOR.

[4] https://www.actronika.com/discover-haptic-technology

Eye tracker

Participants' eye movements were recorded via Tobii Pro Glasses 2 head-mounted Eye Tracking System throughout the experiment. The glasses weigh 45gr and include four eye cameras, a full HD wide-angle scene camera, a gyroscope, and accelerometer sensors (Tobii, 2018). The eye tracker samples gaze position at a rate of 50 Hz with an accuracy of 0.5 degrees of visual angle. The Tobii Pro Lab software (ver. 1.142) was utilised to collect and analyse fixations (>100 ms [168]) on Areas of Interest (AOI). The main output of the data includes videos with gaze positions superimposed.

Questionnaires

After each experimental condition participants completed the following questionnaires:

Nasa-Task Load Index [55] (NASA-TLX). This questionnaire was used for mental workload assessment. It consists of six items with a 0 to 10 Likert scale ("0" represents low and "10" represents high). The six items of mental demand, physical demand, temporal demand, effort, performance, and frustration provide a comprehensive measure of cognitive workload [54].

User Acceptance of New Transport Telematics [146] (UAATT). This questionnaire aims to gather feedback from users on the usefulness and satisfaction with the TOR. The questionnaire includes 9 items rated on a scale from (-2) to (2). See Appendix B.

Single Ease Questionnaire [8] (SEQ). This single-question questionnaire evaluates product usability by asking participants to rate their overall experience on a 1 to 7 scale, indicating ease of use. We used SEQ to assess the ease of detecting the TOR.

4.4 Procedure

Prior to the experiment. After signing consent each participant was presented with a survey (via the online questionnaire platform Qualtrics), which contained demographic questions on age, gender, driving experience, education level and profession.

Next, participants were directed to the simulator and provided with a comprehensive explanation of the experiment's procedure. In a 2-minute training session preceding the experiment, each participant became acquainted with the simulator environment, TORs and NDRTs.

During the experiment. Participants were provided with a Google Pixel 7 smartphone for performing the NDRTs. After putting on and calibrating the Tobii Pro Glasses 2 eye-tracker, they were presented with a pre-recorded driving scenario on the simulator screen, simulating Level 3 automation. Participants were instructed that there was no need to actively monitor the driving environment or maintain control of the vehicle, unless a TOR was issued. They were encouraged to fully engage with the NDRT with their hands off the steering wheel. During the simulation, six TORs were randomly issued either by the IVIS or by the IVIS and the smartphone, depending on the condition. In conditions 1 and 2 the TOR was issued on IVIS only and in conditions 3 and 4 the

TORs were issued by the IVIS and the smartphone simultaneously. See Table 4.1 (participants experienced the different conditions in a counterbalanced order). Participants were asked to press a button on the steering wheel as fast and accurately as possible to mimic taking control after the TOR was issued.

. After the manoeuvre was over, a message saying "Automation On" was displayed on the head-up display and participants were asked to go back to performing the NDRT. In previous literature, the duration of engagement with NDRTs ranged from 60 [15] to 200 seconds [111]. We chose a 120-second average of NDRT engagement before issuing a TOR. This duration is sufficient for engagement with the NDRT while not making the experiment too long. The four conditions were administered in 2 different sessions on separate days to avoid learning bias and fatigue. Each session lasted for approximately one hour.

Post-experiment. Following each of the four conditions, participants were provided with a Samsung Tablet and were given instructions to complete the NASA Task Load Index (NASA-TLX) [55] to assess perceived workload, User Acceptance of New Transport Telematics questionnaires [146] to assess acceptance of the TOR source and Single Ease Questionnaire to asses the ease of detection of the TOR [8]. Questionnaires were administered via Qualtrics. Following the completion of the four conditions, a brief semi-structured interview was held with each participant and the Montreal Cognitive Assessment (MoCA) [109], a cognitive impairment evaluation, was administered by a MoCA certified experimenter to gain a standardised measure of cognitive performance. Participants received £12 at the end of the second session as a compensation for their participation on both sessions of the experiment.

4.5 Results

The results were analysed using R Studio 2022.07.2 Build 576 using the lme4 [10], lmerTest [83] and report [97] Rpackages.

4.5.1 Take-Over Performance

We conducted an ANOVA on a Linear Mixed Model to predict the main effects of Condition (i.e., combination of NDRT and TOR source. See Table 4.1) and Age Group on Reaction Time with a random intercept for each participant and TOR order. All linear model assumptions were met except for normality of residuals. However, according to Schielzeth et al. [130] mixed-effects models are extraordinarily robust, allowing researchers to employ them despite objective violations of the distributional assumptions.

Watching a Video NDRT. Four data points were omitted from the analysis due to instances where participants either dropped the smartphone or failed to adhere to study instructions. We found a significant main effect of Condition with a large effect size

(F(1,10.1)=57.5, p<0.0001; η^2=0.85) and a significant interaction between Condition and Age Group with a medium effect size *(F(1,221.9)=15.6, p<0.0001; η^2=0.07)*. Overall, the (smartphone+IVIS)-TOR led to a significantly lower reaction time *(Estimated Mean (EM)=2.06 sec, SE=0.2 sec)* than issuing the TOR only on the IVIS *(EM=2.65 sec, SE=0.2 sec)*. *Post hoc* analyses with a Bonferroni adjustment revealed that while reaction times of both groups were lower in the (smartphone+IVIS)-TOR condition compared to IVIS-TOR, the 70+ age group exhibited a significantly greater reduction of reaction time in the (smartphone+IVIS)-TOR condition *(ΔEM= 0.93 sec, p<0.0001)* compared to the 60-69 group *(ΔEM= 0.26 sec, p<0.05)*. See Figure 4.5.

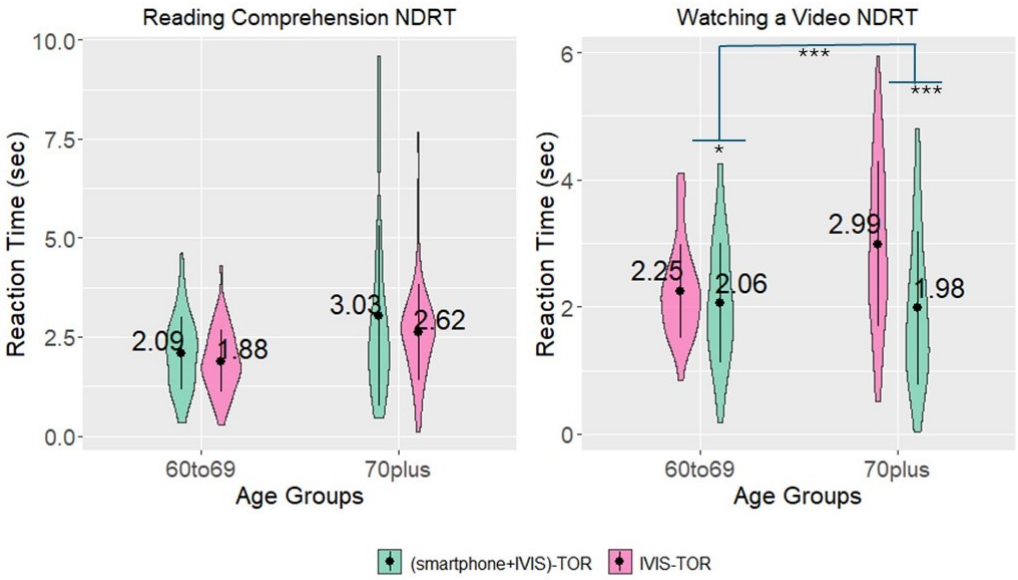


Figure 4.5: The interaction between Age Group and Condition for each NDRT for the dependent variable Reaction Time. The dots represent mean values. *** - P<0.001, ** - P<0.01, *- P<0.05.

Reading Comprehension NDRT. For the Reading Comprehension NDRT the fitted model did not yield any significant effects.

4.5.2 NASA-TLX Overall Workload

After checking that model assumptions were met, we conducted an ANOVA on a Linear Mixed Model to predict the main effects of Condition and Age Group on NASA-TLX scores with a random intercept for each participant.

Watching a Video NDRT. We found a significant interaction between Condition and Age group *(F(1,20)=5.2, p<0.05; η^2=0.21)*. *Post hoc* analyses with a Bonferroni adjustment revealed that for the age group 60-69 there was no significant difference in the overall NASA-TLX scores between conditions (*p=0.13*). The 70+ group exhibited a significantly higher (*p<0.05*) overall NASA-TLX score for the IVIS-TOR *(Estimated Mean (EM)=18,*

SE=2.49) than the (smartphone+IVIS)-TOR condition *(EM=12.9, SE=2.49)*. See Figure 4.6. The only item result that was significant was the temporal demand. The IVIS-TOR showed a significantly higher temporal demand score for the 70+ group compared to the (smartphone+IVIS)-TOR condition ($p<0.05$).

Reading Comprehension NDRT. We found a significant interaction between Condition and Age group *($F(1,20)=8.1$, $p<0.01$; $\eta^2=0.29$)*. *Post hoc* analyses with a Bonferroni adjustment revealed that, for the age group 60-69, there was no significant difference in the overall NASA-TLX scores between conditions ($p=0.37$). The 70+ group however, exhibited a significantly higher ($p<0.001$) overall NASA-TLX score for the (smartphone+IVIS)-TOR condition *(EM=28.8, SE=2.84)* over the IVIS-TOR *(EM=18.1, SE=2.84)*. See Figure 4.6. The only item result that was significant was the Effort. The (smartphone+IVIS)-TOR led to a significantly higher effort score for the 70+ group compared to the IVIS-TOR condition ($p<0.001$).

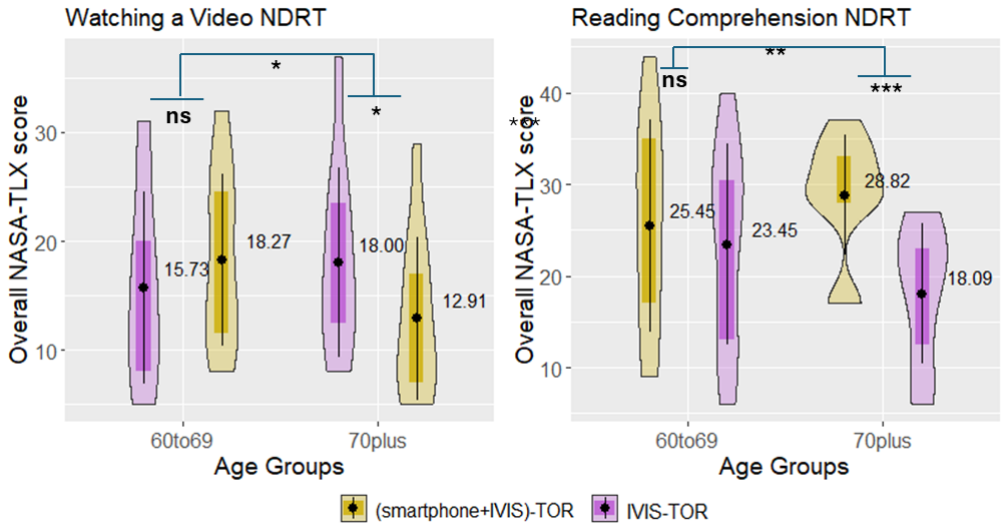


Figure 4.6: The interaction between Age Group and Condition for each NDRT for the dependent variable Overall NASA-TLX score. The dots represent mean values. *** - P<0.001, ** - P<0.01, *- P<0.05.

4.5.3 User Acceptance of Advanced Transport Telematics Questionnaire (UAATT)

A reliability check yielded a sufficiently high Cronbach's alpha of 0.81 [146]. Items 1,3,5,7,9 were averaged for the Usefulness score and items 2,4,6,8 were averaged to get the Satisfaction score [146]. See Figure 4.7 for the items. The data did not have a normal distribution (via the Shapiro-Wilk test), so we conducted an Aligned-Rank Transform (ART) one-way ANOVA exploring the main effects of Condition and Age Group with a random intercept for each participant on the dependent variables Satisfaction and Usefulness.

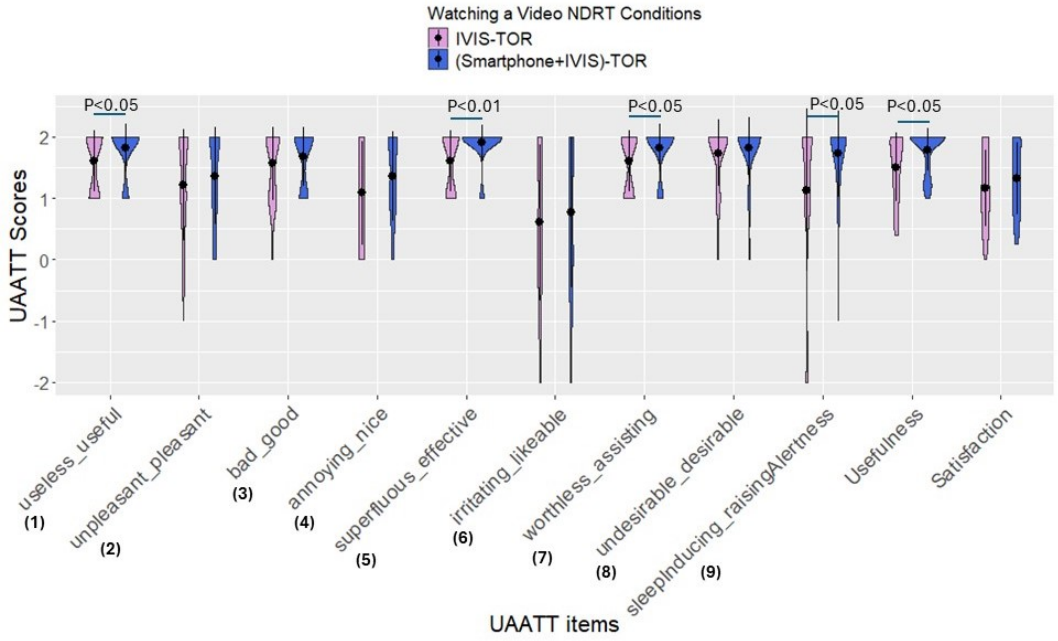


Figure 4.7: UAATT Questionnaire scores of Usefulness and Satisfaction for watching a video NDRT. On the x-axis are the items of the questionnaire. "Usefulness" and "Satisfaction" represent the averaged values.

.

Watching a Video NDRT. For the dependent variable Usefulness, the main effect of Condition was statistically significant *($F(1,21.4) = 9.38$, $p<0.01$; $\eta^2 = 0.3$)*. The (smartphone+IVIS)-TOR condition yielded significantly higher Usefulness scores *($M=1.79$, $SD=0.35$)* than the IVIS-TOR *($M=1.5$, $SD=0.56$)*. See Figure 4.7. The interaction between Age group and Condition was not significant *($F(1,21.4) = 2.3$, $p = 0.14$)*. For Satisfaction, the model did not yield any significant main effects. See Figure 4.7.

Reading Comprehension NDRT. For Satisfaction and Usefulness both models did not yield any significant main effects.

4.5.4 Single Ease Questionnaire

The data did not have a normal distribution (via the Shapiro-Wilk test), so we conducted an Aligned-Rank Transform (ART) one-way ANOVA exploring the main effects of Condition and Age Group with a random intercept for each participant on the scores of Ease of Detection of the TOR.

Watching a Video NDRT. We found a significant main effect of Condition *($F(1,21.4) = 5.98$, $p = 0.022$; $\eta^2 = 0.2$)*. The (smartphone+IVIS)-TOR condition yielded significantly higher Ease of Detection scores *($M=6.73$, $SD=0.59$)* than IVIS-TOR *($M=5.76$, $SD=1.61$)*. See Figure 4.8. The interaction between Age group and Condition was not significant *($F(1,22.1) = 0.97$, $p = 0.33$)*.

Reading Comprehension NDRT. The model did not yield any significant effects. See Figure 4.8.

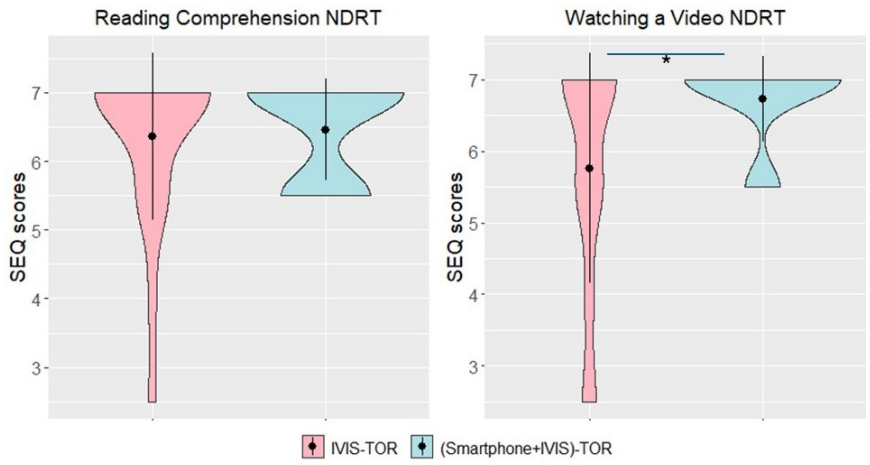


Figure 4.8: Single Ease Questionnaire scores for each NDRT measuring ease of detection of the TOR (1= hard to detect, 7= easy to detect). *-P<0.05.

4.5.5 Hazard Perception

The initial model in each analysis included the following fixed effects: Age Group and Condition. The random effects of the initial model included participant and TOR order. The final model of each analysis was achieved via a backward elimination procedure where all non-significant interaction effects were removed from the model. For significant fixed effects with more than two levels, post hoc pairwise contrasts comparisons analysis was applied with Bonferroni adjustment.

Watching a Video NDRT

Number of Fixations (>100 ms) on AOI The dependent variable Number of Fixations on the AOI had a Poisson distribution. We conducted an ANOVA on a Poisson regression model with a log link function within the Generalised Linear Mixed Models (GLMM) framework.

The model's total explanatory power was moderate (conditional R2 = 0.22) [5]. Within this model a significant main effect of Condition (χ^2 = *3.93, df = 1, p<0.05*) was found, (smartphone+IVIS)-TOR yielded a higher number of fixations on the AOI (*Estimated Mean (EM)=4.49, SE=0.52*) than IVIS-TOR condition (*EM=3.19, SE=0.41*). Age Group did not show a significant effect on the number of fixations (χ^2 = *0.17, df=1, p=0.67*). No interaction between Age Group and Condition was found (χ^2 = *1.23, df = 1, p=0.26*).

[5]Whilst it is difficult to calculate the actual R2 value for generalised linear models, the theoretical value was calculated using the MuMIn package [9]. R2 values above 0.2 are generally considered good [101].

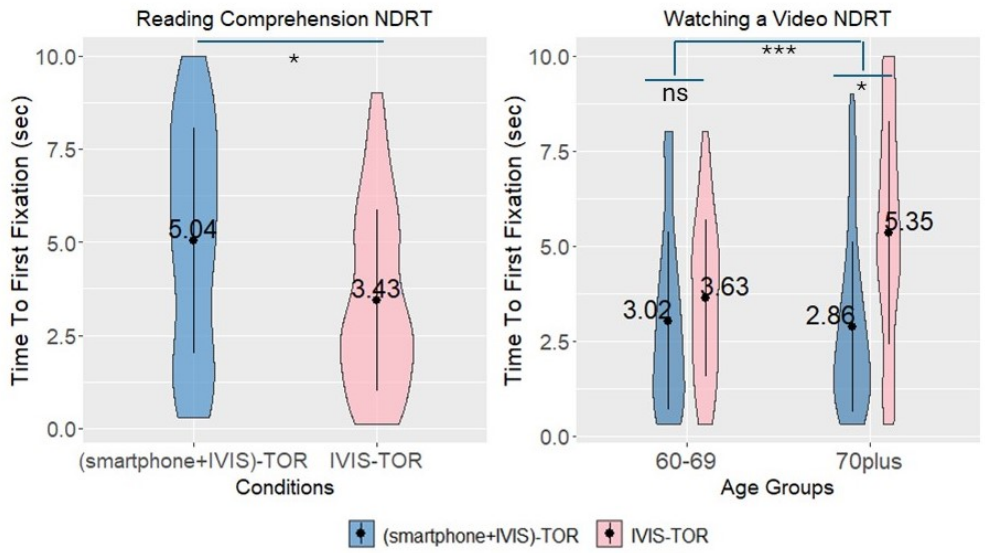


Figure 4.9: Time To First Fixation on the AOI for each NDRT. The dots represent the mean values. ***-P<0.001, *-P<0.05.

Time to First Fixation on AOI (TTFF) We conducted an ANOVA on a linear regression model within the Linear Mixed Models (LMM) framework with a normal distribution for the dependent variable. Normality of the data was confirmed via Shapiro-Wilk test.

The interaction between Age Group and Condition was significant (χ^2 = *11.81, df = 1, p<0.001,* η^2 *=0.06*). *Post hoc* analyses with a Bonferroni adjustment revealed that for participants in the age group 60-69, no significant difference was observed between conditions (*p=0.33*). However, in the age group 70+, a significant difference emerged between conditions (*p<0.05*). The 70+ group had significantly lower TTFF when the TOR was (smartphone+IVIS)-TOR (*Estimated Mean=2.83 sec, SE=0.84 sec*) than when the TOR was on the IVIS only (*EM=5.81 sec, SE=0.88 sec*). See Figure 4.9.

Hazard Detection For the dependent variable Hazard Detection, that has a binomial distribution, we conducted an ANOVA on a logistic regression model with a logit link function within the Generalised Linear Mixed Models (GLMM) framework. Hazard Detection was treated as a binary variable that describes whether a participant had at least one fixation of at least 100 ms at the AOI (coded as "1") or not (coded as "0").

The model's total explanatory power was moderate (conditional R2 = 0.32). We found a significant interaction between Condition and Age Group (χ^2 = *4.46, df = 1, p<0.05*). *Post hoc* analyses with a Bonferroni adjustment revealed that, for the age group 60-69, there was no significant difference in the estimated probability to detect the hazard between conditions (*p=0.8*). The 70+ group however, exhibited a significantly higher estimated probability to detect the hazard for the (smartphone+IVIS)-TOR condition (*Estimated Mean (EM)= 0.97, SE=0.022, p<0.01*) over the IVIS-TOR *(EM=0.82, SE=0.08)*.

Reading Comprehension NDRT

Number of Fixations (>100 ms) on AOI The dependent variable Number of Fixations on the AOI has Poisson distribution. We conducted an ANOVA on a Poisson regression model with a log link function within the Generalised Linear Mixed Models (GLMM) framework.

The model's total explanatory power was moderate (conditional R2 = 0.5). Within this model a significant main effect of Age Group (χ^2 = *8.51, df = 1, p<0.01*) was found. The 60-69 group had a significantly larger number of fixations on the hazard (*EM=3.98, SE=0.34*) than the 70+ group (*Estimated Mean (EM)=2.76, SE=0.28*). No interaction was found between Age Group and Condition (χ^2 = *0.15, df = 1, p=0.6*).

Time to First Fixation on AOI (TTFF) We conducted an ANOVA on a linear regression model within the Linear Mixed Models (LMM) framework with a normal distribution for the dependent variable. Normality of the data was confirmed via Shapiro-Wilk test.

The model revealed a significant main effect of Condition (χ^2 = *4.26, df = 1, p<0.05,* η^2 *=0.26*). The IVIS-TOR condition led to a significantly lower TTFF (*EM=3.33 sec, SE=0.62 sec*) compared to (smartphone+IVIS)-TOR condition (*EM=4.94 sec, SE=0.63 sec*). See Figure 4.9. No significant interaction was found between Age Group and Condition.

Hazard Detection For the dependent variable Hazard Detection, that has a binomial distribution, we conducted an ANOVA on a logistic regression model with a logit link function within the Generalized Linear Mixed Models (GLMM) framework. Hazard Detection was treated as a binary variable that describes whether a participant had at least one fixation of at least 100 ms at the AOI (coded as "1") or not (coded as "0").

The model did not yield any significant effects.

4.5.6 Post Experiment Interviews

Following the guidance of Braun et al. [17] for thematic analysis, 32 unique codes were extracted from the data using NVivo (https://lumivero.com/ products/nvivo/). The authors sorted these through an iterative process into three themes based on code similarity. We present and explain the themes supported by quotes from participants.

Theme 1: Visual message location. Since the visual component of the TOR was presented at the top of the smartphone screen (See Figure 6.1), participants were tempted to continue performing the task. Par6 (71 yo) said, *"Your first reaction is to ignore it [Visual message on the top of smartphone screen] and continue with what we're doing. But then of course you can't do that so you stop doing the task"*.

They suggested that the visual message of the TOR appear in the middle of the screen blocking the task. For example, Par13 (62yo) said, *"If the message was in the middle it will*

block the view so I will have no choice but be immediately attentive to the road".

Theme 2: Task switching. Reclaiming manual control after a TOR was harder for the reading task than the video for all but two participants. Participants had to make the decision on whether to finish the statement or notion they were reading before disengaging from the task. Also, returning back to reading was harder than returning to the video since participants had to revisit the last sentence that they read. Par10 (69yo) said, *"Was a bit more of a challenge because I had to remember where I had got to and or what the question was. So I felt there was maybe a revisiting the text, whereas with the video, I just picked up the video"*.

Theme 3: TOR source. Duality in TOR presentation (i.e., presenting on both the smartphone and IVIS) was preferred by participants. (smartphone+IVIS)-TOR was thought to be effective in shifting attention away from the NDRT.

Par6 (71yo) explained, *"I can see the absolute necessity in having it [on the smartphone]. Because if you get too engrossed in your device, how do you switch off to something that is external to it [..] I would especially be the favour of having the warning coming on whatever device you are using"*. Par12 (67yo) elaborated, *"Not having the phone alert was actually quite frustrating, mentally taxing and made you feel less confident"*.

4.6 Discussion

The goal of this study was twofold. Firstly, we sought to establish how the TOR source ((smartphone+IVIS) vs. IVIS) and the NDRT cognitive demand affect older drivers' reaction times, hazard perception and user acceptance. Secondly, we investigated whether these effects differ for older drivers as a heterogeneous age group.

- **RQ1** How does a (smartphone+IVIS)-TOR affect takeover reaction time compared to IVIS-TOR across two cohorts of older drivers, considering two NDRTs with varying levels of cognitive demand?

Watching a Video (Lower cognitive demand NDRT). Our results indicate that a TOR presented on a smartphone and the IVIS leads to a significantly lower reaction time for both groups of older drivers, compared to just presenting it on the IVIS. While the reduction in reaction time was significant for both age groups, our findings show that drivers aged 70+ benefit more as their RT was reduced by 1.01 seconds compared to 0.19 seconds for the 60-69 group. Previous research has found that older drivers (mean age = 64.3) experience impaired takeover performance after a visual-auditory TOR, particularly when switching between multiple NDRTs [158]. Drivers aged 70 and older were slower to respond to and complete takeover requests than those aged 60-69 [93]. The results from this study suggest that these reaction times can be improved by integrating the TOR into the device where the NDRT is presented, in this case a smartphone held by the driver.

There was a significant interaction between the TOR presentation source ((smartphone+IVIS) vs. IVIS) and age group on the overall NASA- TLX scores. Participants aged 70+ had significantly higher overall workload and temporal demand scores in the IVIS-TOR condition when compared to the (smartphone+IVIS)-TOR condition. This suggests that adding a TOR to the smartphone could reduce cognitive demand for 70+ drivers, important for when they need to drive the vehicle. Previous research highlights challenges older individuals face when dividing attention across space [77]. For the 70+ group, displaying a TOR on a smartphone may help distribute cognitive load more efficiently, reducing temporal demand. The Ease of Detection Questionnaire results showed that the (smartphone+IVIS)-TOR was easier to detect than the IVIS-TOR. In practical terms, incorporating the TOR into the smartphone can significantly enhance the responsiveness of older drivers to takeover requests in Level 3 cars, contributing to safer and more efficient transitions between automated and manual driving modes.

Reading Comprehension (Higher cognitive demand NDRT). For the more cognitively demanding NDRT, there was no significant difference in reaction time between (smartphone+IVIS)-TOR and IVIS-TOR. Considering that the cognitive demand imposed by the reading task is inherently high [108], participants might have consistently allocated a substantial amount of cognitive resources to the task, minimising the observable differences in reaction times across the TOR presentation modes. The cognitive demand of the NDRT has been identified as a factor influencing reaction times to TOR [108]. The absence of a significant difference in overall NASA-TLX scores for the 60-69 age group implies a consistent cognitive workload allocation across both TOR sources for this group. Consequently, no significant impact on reaction times was observed across the TOR presentation sources for this age group. However, the 70+ age group showed a notable difference, experiencing a significantly higher overall NASA-TLX score for the (smartphone+IVIS)-TOR condition compared to the IVIS-TOR condition. This discrepancy was particularly pronounced in the Effort item of the questionnaire. This may be attributed to the additional mental effort required to process the TOR information. The greater impact on the 70+ age group may be attributed to their relatively lower cognitive adaptability, making managing higher task workloads less effective than their younger counterparts [112]. The 70+ group reaction times were slightly higher for the (smartphone+IVIS)-TOR presentation source compared to IVIS-TOR, however the difference did not reach significance.

- **RQ2** How does (smartphone+IVIS)-TOR affect hazard perception compared to IVIS-TOR across two cohorts of older drivers, considering two NDRTs with varying levels of cognitive demand?

Watching a Video (Lower cognitive demand NDRT). The eye tracking data revealed that the (smartphone+IVIS)-TOR condition led to a higher number of fixations on hazards compared to the IVIS-TOR condition, irrespective of age. A commonly employed metric for evaluating hazard perception is the number of fixations on the hazard; the

greater the number of fixations on the hazard, the greater the likelihood that the driver perceived it [168, 16]. Displaying a TOR on a smartphone and the IVIS improved hazard perception, helping older drivers make informed decisions upon regaining control, thereby enhancing overall safety.

For the 60-69 age group, no significant difference in time to first fixation on the hazard was observed between the (smartphone+IVIS)-TOR and IVIS-TOR conditions. This suggests that this age group might have sufficient cognitive resources to shift their attention even without the additional support from the smartphone. In contrast, the 70+ age group exhibited significantly shorter TTFF when the TOR was on the smartphone and IVIS compared to the IVIS alone. The higher perceived workload in the IVIS-TOR condition, as indicated by overall NASA-TLX results, could explain this difference. Presenting the TOR on the smartphone and IVIS may have alleviated cognitive load for the 70+ group, enabling quicker attention allocation to hazards. This age-specific impact holds promise in light of studies like Horswill et al. [61], which highlight that age-related declines in hazard perception exacerbate in drivers aged 75 and above, increasing crash risk. Our findings suggest that presenting the TOR on the NDRT device shows potential in compensating for the hazard perception decline in older drivers.

The (smartphone+IVIS)-TOR condition resulted in a higher probability of hazard detection for the 70+ age group compared to the IVIS-TOR condition. Presenting the reason for the take-over request on the smartphone, both visually and aurally, likely lowered cognitive load, as confirmed by NASA-TLX results. This prompted a quicker shift of attention to the hazard and increased the likelihood of detecting it within the TOR time budget. The 60-69 age group, however, might have had sufficient cognitive resources to shift their attention and detect the hazard even without the additional support from the smartphone.

Reading Comprehension (Higher cognitive demand NDRT). Irrespective of age group, there was no significant effect of the TOR presentation source on the number of fixations and the (smartphone+IVIS)-TOR presentation source led to longer TTFF. NASA- TLX results showed that (smartphone+IVIS)-TOR did not lower the perceived workload compared to the IVIS-TOR, instead it led to higher workload for the 70+ group. Subjective workload measured by NASA-TLX has high correlation with gaze metrics [38]. For example, higher task complexity leads to shorter duration of fixations on the hazard [38]. Therefore, we speculate that drivers may have experienced challenges in managing the high workload of both the reading task and the (smartphone+IVIS)-TOR, which resulted in reduced cognitive resources available for shifting their attention to the hazard and perceiving it.

There was no significant difference between the age groups or the conditions in relation to hazard detection. The lack of significant differences in hazard detection across age groups may be attributed to the informative nature of the TOR. The TOR informed participants of what the hazard was and they demonstrated a comparable ability to detect it within the given TOR duration.

- **RQ3** Does (smartphone+IVIS)-TOR enhance TOR acceptance and detection compared to IVIS-TOR across two cohorts of older drivers, considering two NDRTs with varying levels of cognitive demand?

Watching a Video (Lower cognitive demand NDRT). Participants found the (smartphone+IVIS)-TOR more useful than the IVIS-TOR. The favourable impact of the (smartphone+IVIS)-TOR was corroborated in Theme 3 of the post-experiment interviews, where the majority of participants stated that providing the TOR on a smartphone could boost safety because the TOR is triggered on the device to which their attention is directed. Satisfaction scores were the same across presentation sources. Satisfaction was linked to TOR likeability, suggesting that the (smartphone+IVIS)-TOR design should be further improved.

Reading Comprehension (Higher cognitive demand NDRT). There was no significant effect of TOR presentation source on neither Satisfaction nor Usefulness. The high demand of the visual cognitive task may have consumed participants' attention and cognitive resources. In such cases, the source of TOR presentation may not have had a significant impact on satisfaction or perceived usefulness because participants were primarily focused on the demanding cognitive task

4.6.1 Summary

Most researchers reported that engagement with NDRTs has negatively influenced the takeover performance [80] and hazard perception ability [40] of older drivers. To ensure Level 3 cars are safe for this age group, addressing the challenges of regaining control after NDRT engagement is crucial. In this study, we proposed a (smartphone+IVIS)-TOR as a solution to mitigate this challenge. The findings highlight an intricate relationship between the proposed solution, (smartphone+IVIS)-TOR, and task cognitive demand. For NDRTs with lower cognitive demand, smartphone integrated TORs can be a valuable asset for older drivers. Linking back to the deterioration in executive functions, a (smartphone+IVIS)-TOR can reduce driving-related EF deficiencies in "Shifting" and "Inhibition". Presenting the TOR on the NDRT-device, which is already in the drivers' visual and cognitive focus, shortened reaction times for both age groups and led to more fixations on the hazard. Also, it increased hazard detection probability and reduced time to first fixation for 70+ drivers, improving their ability to shift the visual attention to the hazard. The subjective evaluation of the (smartphone+IVIS)-TOR reveals that it is more useful than IVIS-TOR.

Regarding the NDRT with higher cognitive demand, the solution of (smartphone+IVIS)-TOR was not of benefit for older drivers. It inversely led to longer time to first fixation for both age groups. According to the multiple resource theory of Wickens et al. [59], cognitive resources are limited, and when a task requires high cognitive demand, there is competition for these resources. The cognitive resources may have been heavily taxed by the reading task that when a TOR was presented on the smartphone, it required additional cognitive resources to process it, leading to competition for the already scarce

cognitive resources.

Therefore, further design iterations are required to bring the (smartphone+IVIS)-TOR to a point where it is beneficial to older drivers when engaged with an NDRT with high cognitive demand. As reported by the participants in the post-experiment interviews, they preferred duality in TOR presentation. They favoured, nevertheless, a TOR that visually obstructs the task. In light of this, subsequent iterations of (smartphone+IVIS)-TORs ought to obstruct the NDRT visually. Then, an evaluation of the new design should be conducted for both NDRT cognitive demand levels.

4.7 Limitations and Future Work

The NDRTs used in the study are of the types that users said they would engage with while in a Level 3 vehicle [92, 141]. The reading task represents a visual-cognitive task with higher cognitive demand and watching a video represents a visual- cognitive task with lower cognitive demand. Yet they are not entirely representative of NDRTs as a whole. Results from visual-cognitive tasks may not generalise to tasks that use different resources (e.g., physical-auditory tasks) [164]. Therefore, future research should expand the current knowledge on TOR design and presentation to include a wider variety of NDRTs.

The majority of our participants (≈70% of the 60-69 group and ≈60% of the 70+ group) had at least a bachelor's degree. Stern et al. [143] suggests that individuals with higher levels of education may exhibit better cognitive abilities. Therefore, our older driver's cohort may not be representative of the general older driver's population. To ensure that Level 3 cars are of benefit to the general population of older drivers, a greater range of driver demographics and a larger sample size should be included in future studies.

Whilst the road footage used in the study was collected from a wide range of real road types and environments in the UK, the majority of the hazards occurred in urban or suburban roads. This is consistent with areas where most road collisions occur [49]. However, due to the limited experiment duration, TORs were issued at a relatively high frequency (on average one TOR every 2 minutes) than would occur in real driving. The higher frequency of TORs may have caused drivers to be more vigilant than they would be in a more realistic setting with longer gaps between TORs which could have influenced their reaction times and hazard perception.

Hazard criticality can range between unexpected road events (higher criticality) to pre-planned take overs (lower criticality). In this study the time budget of TORs was fixed (10 seconds) regardless of the type of hazard. Future work could leverage this design and develop TORs with varying time budgets depending on the criticality of the hazardous event.

The TOR in its current design brought major benefits to older drivers while performing an NDRT with low visual-cognitive demand. However, it did not improve performance, hazard perception or user experience when the NDRT was of high visual-

cognitive demand. Consequently, a more nuanced examination of the TOR design is needed.

4.8 Conclusion

Older adults may face challenges going back into the driving loop after a take-over request due to deterioration in cognitive and physical abilities [99, 87, 62]. These challenges intensify when drivers' attention is allocated to the NDRT device but the TOR is presented by the in-vehicle information system alone, ignoring the NDRT-device. Hence, investigating how the source of TOR affects control transition for older drivers is crucial.

Our study explored the impact of TOR presentation source and NDRT cognitive demand on older drivers, focusing on reaction times, hazard perception, and user acceptance of the TOR. We triggered the TOR on a smartphone where the NDRT was displayed (i.e., where the driver's attention was allocated) along with the in-vehicle information system ((smartphone+IVIS)-TOR), and compared it to the TOR displayed only on the IVIS (IVIS-TOR). We tested two NDRTs with different levels of cognitive demand to investigate how this affected take over performance, hazard perception and acceptance of the TOR. For the NDRT with lower cognitive demand (watching a video), a (smartphone+IVIS)-TOR significantly reduced reaction times for older drivers; a greater reduction in RT was observed for drivers aged 70+. Eye-tracking data revealed that the (smartphone+IVIS)-TOR condition led to increased number of fixations on hazards for both age groups, a shorter time to first fixation and a higher probability of hazard detection for the 70+ group, improving hazard perception. Furthermore, (smartphone+IVIS)-TOR was perceived as more useful and easier to detect than IVIS-TOR for the lower cognitive demand NDRT. For the higher cognitive demand NDRT, no benefits were found from the dual presentation of the TOR implying that further design refinements should be made. According to Srour-Zreik et al. [141], older drivers will initially use NDRTs with low cognitive demand until they gain trust in Level 3. This underscores the practical relevance of our findings, as the advantages conferred by (smartphone+IVIS)-TOR are particularly noteworthy during early stages of Level 3 deployment. This study takes the first steps towards evaluating TOR incorporation into the NDRT device for older drivers. Our findings should be useful for practitioners and designers involved in the design and development of TOR features in Level 3 cars.

Chapter 5

All-inclusive TORs: A Cross-Cultural and Age-Sensitive Design for Take-Over Requests in Level 3 Cars (CHI '25)

5.1 Introduction

Autonomous vehicles hold the promise of transforming transportation, offering enhanced safety, energy efficiency, and accessibility while reshaping how we use urban space and manage transportation costs [12]. Among the levels of driving automation defined by the Society of Automotive Engineers (SAE) (from Level 0: manual driving to Level 5: full automation) [1], Level 3 vehicles represent a critical milestone. They allow drivers to disengage from the act of driving and immerse themselves in other activities referred to as Non-Driving-Related Tasks (NDRTs) [1]. However, drivers must remain vigilant, ready to assume manual control when a Take-Over Request (TOR) is issued if the automated system encounters a situation it cannot handle. The introduction of Level 3 cars thus brings a complex and unresolved multi-disciplinary problem: as drivers become increasingly detached from the driving task, their response times slow and their situational awareness diminishes, particularly when engaged in visually demanding NDRTs [43, 168]. With Level 3 vehicles now available in markets such as the US and Germany [53], there is an urgent need to rethink and optimise TOR design to mitigate the challenge of control transition.

Effective TOR design requires a deep understanding of the factors that influence driver behaviour during these control handovers. Age, culture, and NDRT type are key elements that shape how drivers perceive and react to TORs. For example, Level 3 vehicles have the potential to extend the mobility of older drivers [7], enhancing their independence and quality of life [141]. Yet, older drivers often struggle with control transitions due to executive function decline [24], making it harder for them to respond

quickly and accurately to TORs [117, 93]. Additionally, cultural differences shape driving behaviour [12, 139, 148] - what works in one cultural context may be less effective in another. The cognitive and sensory demands of different NDRTs further complicate these transitions, significantly impacting the quality of a take-over response [168, 165]. Yet, despite the clear impact of age, NDRT type and culture on driver behaviour, current TOR designs tend to follow a one-size-fits-all approach, missing the nuanced cognitive and sensory needs that vary across age groups and cultural backgrounds. Critically, these designs often neglect how these factors interact and amplify one another, resulting in TORs that are poorly suited to the diverse profiles of drivers. For instance, older drivers might require simpler and more disruptive alerts to help them shift from non-driving activities back to active driving. Likewise, drivers from different cultural backgrounds may perceive hazards differently [148], necessitating tailored design strategies. This lack of tailored design can lead to delayed, inadequate, or entirely missed responses to TORs, ultimately compromising safety during control transitions in Level 3 cars.

Using participatory design methods in a real car setting, this study is the first to explore how the interplay between age and cultural factors impacts TOR design requirements across two distinct cultures: the UK representing high compliance with traffic laws; and Israel, characterised by lower compliance, and across two age groups (60+ and 22-32 year-old drivers). Participants, grouped by age and culture, were seated in the front seats of the car where they engaged in NDRTs and designed TORs. They used stationery materials, leveraging both the car's interior and the NDRT devices to facilitate their designs. We performed a thematic analysis of the design session outcomes to identify the features that TORs should include to facilitate control transitions in Level 3 autonomous driving, over four different NDRTs (Reading, Texting, Watching a video and Chatting with a passenger). From this analysis, we created a taxonomy of TOR features and, for the first time, synthesised versatile TOR concepts tailored to the needs of older and younger drivers across the two cultures. The results revealed both distinct and similar preferences across groups: for example, all drivers preferred the repurposing of the NDRT device to issue the TOR. Older drivers preferred tri-modal TORs and favoured NDRT suspension, whereas younger drivers opted for bi-modal TORs, maintaining control over NDRT interruptions. Influenced by regional conflict and lower traffic law compliance, Israeli participants proposed a Risk-Aware TOR that assesses hazard severity. The practical implications of this study include providing guidelines for designing take-over requests that consider diverse types of NDRTs and cater for the requirements of two age groups across two cultures.
We contribute:

- The first cross-cultural participatory design study on TORs- with participants designing in a real car setting;
- A novel approach for designing interfaces that utilise the NDRT device as a source for TOR issuance;
- A taxonomy of TOR interface components including messages, their placements,

and modalities;

- Identification of themes representing participant expectations of TOR interfaces, highlighting differing requirements and perspectives across two age groups and two cultures.

5.2 Related Work

5.2.1 The Challenge of Control Transition in Level 3 Cars

Research consistently shows that several factors influence how quickly drivers can regain control after a TOR, including the driver's experience, the TOR design and their engagement in secondary tasks [170]. Visual-manual tasks (i.e., an activity that requires both visual attention and manual interaction such as texting), in particular, are known to significantly impair driver performance and increase take-over time [104, 151, 124]. A simulator study found that visual-manual NDRTs lower road hazard awareness and increase cognitive workload, as reflected in eye-tracking and NASA-TLX scores [168]. Likewise, cognitively demanding tasks have a greater impact in critical situations, while motoric distractions are more disruptive in less urgent scenarios [43]. Research indicates that different types of NDRTs impact take-over reaction time differently. Smartphone use causes greater delays than watching a video [165], while high-workload tasks like reading or texting lead to longer reaction times [108]. These studies underscore the critical need to address the interplay between TOR and NDRT types, emphasising the importance of designing TORs that are tailored to specific NDRT activities. In response to this need, our study conducted participatory design sessions aimed at developing TORs customised for different NDRT types.

The challenge of control transitions in Level 3 autonomous vehicles is particularly pronounced for older drivers, who may experience age-related declines in cognitive, motor and sensory functions [117, 31, 158]. Declines in cognitive flexibility, such as reduced ability to shift and distribute visual attention [7, 116, 77, 117], may slow their response to takeover requests when transitioning from NDRTs. Sensory impairments, including diminished peripheral vision and contrast sensitivity [7, 116, 77, 117], may impair hazard detection, further increasing risk. Motor declines, such as reduced muscle strength and joint flexibility [7, 117, 77], may make quick reactions difficult in critical situations. These declines may impair older drivers' ability to respond effectively to driving alerts after engaging in NDRTs [117], particularly when switching between multiple NDRTs [158] or engaging with cognitively demanding NDRTs [31, 140]. When compared with their younger counterparts, older drivers exhibit slower reaction times [91], shorter mirror fixation duration [116] and deteriorating hazard perception ability [37, 60, 7]. Recognising their cognitive and physical decline, older drivers have advocated for the interruption of ongoing NDRTs by the takeover request on the NDRT device to ease the transition back to manual control [141, 92]. Although the referenced studies [116, 31, 91, 158] demonstrate that the negative impacts of NDRT engagement are more pronounced for older

drivers (e.g., poorer take-over performance), they remain significantly underrepresented in research focused on user interface design for control transitions. In a literature review, Happee et al. [82] examined 25 studies that specifically addressed the effectiveness of user interfaces during control transitions in Level 3 vehicles, but none of these studies included older drivers, instead they focused on younger or middle-age drivers. This makes it imperative to develop TORs customised for older drivers, carefully considering the demands of NDRTs, to assist them in smoothly and safely transitioning from non-driving activities to driving.

5.2.2 Take-Over Request (TOR) Design

The effectiveness of a TOR is determined by four key elements: presentation modality, timing of the alert, reliability of the alert, and the required level of cognitive processing [71]. Uni-modal TORs have proven to be less effective than multi-modal ones, with each modality offering specific advantages: haptic feedback prompts immediate reactions, while auditory cues provide more stable responses [166]. Multi-modal alerts, combining visual, aural, and vibrotactile cues, generally lead to quicker responses among younger drivers [120, 64], but can also increase feelings of urgency [120], annoyance [120] and increase workload [166]. Beyond modality choice, studies have also explored other aspects of TOR design. For instance, a TOR allowing drivers to monitor the vehicle's performance for two seconds before taking control slightly reduced takeover time by 0.3 seconds [42]. Another study found that a fixed Human-Machine Interface (HMI) design, which transfers control after a set duration, led to better takeover performance than a variable HMI that responded to drivers' visual attention shifts [105]. While these studies have made strides in understanding TOR design effectiveness, gaps remain in exploring *how* these TOR interfaces are designed and whether they consider the preferences and needs of diverse user groups and NDRT types [82, 92]. Another aspect explored in the literature is the source from which the TOR is issued (i.e, vehicle information system, personal device). A simulator study by Melcher et al. [103] found that including the visual component of the TOR on the driver's mobile phone increased TOR acceptance and perceived safety without affecting response times. Similarly, Wintersberger et al. [157] showed that a TOR presented on the NDRT device led to reduced stress, increased trustworthiness, and a tendency for enhanced takeover performance. These studies overlooked the cognitive demands of different NDRTs and lacked clarity on the rationale behind the TOR design. Thus far, research on control transitions for younger drivers has exposed critical gaps. First, studies that address the cognitive demands of various NDRTs in shaping design requirements are scarce. Second, no research has incorporated the NDRT device itself into the design process—an essential step given younger drivers' strong willingness to engage with device-based NDRTs in Level 3 vehicles [141, 20]. Third, research comparing TOR design requirements of younger drivers from different driving cultures is absent. Therefore, involving younger drivers in participatory design sessions that consider multiple NDRTs, incorporate NDRT devices, and explore diverse driving cultures

is crucial for the development of effective TORs.

Limited research on TOR design has addressed older drivers' needs. [64] found that, similar to younger drivers, older drivers detect tri-modal TORs more easily and react faster compared to uni-modal or bi-modal ones. For auditory components, older drivers prefer speech-based over non-speech-based stimuli, with speech providing more detailed information [128]. Additionally, tactile components in alerts, such as auditory-tactile or visual-auditory-tactile combinations, significantly reduce brake response times for older drivers [119]. However, current TOR designs have not been optimized to coordinate effectively with different NDRTs, which are known to affect control transitions—especially for older adults, whose cognitive abilities may vary [117, 158, 91]. Additionally, Li et al. [92] highlight the need to design and evaluate TORs considering their source, particularly when issued on the device used for performing the NDRT. Srour-Zreik et al. [140] found that redundancy in TORs, issuing them from both the NDRT-device and HUD (Head-Up Display: the projection of driving info onto windscreen), significantly improved performance in control transitions for older drivers during low visual-cognitive demand tasks but not during high-demand tasks. Therefore, the effectiveness of redundant TORs may depend on the cognitive load of the NDRT, suggesting that a more targeted approach to TOR design is necessary [140, 92].

5.2.3 Cross-cultural studies on driving behaviour in Level 3 cars

Culture is known to significantly influence driving behaviour and compliance with traffic laws [173, 152, 46, 148, 27]. For example, Wang et el. [152] investigated risk-avoiding behaviours between Chinese and German drivers in scenarios involving vulnerable traffic participants, revealing significant cultural differences in how drivers respond to potential risks. In two of the three driving scenarios, Chinese participants had significantly longer take-over time than German participants which may be attributed to different threshold in risk perception. In developing countries, the driving environment is more complex, requiring different driving behaviours [148]. For example, Malaysian roads exhibit higher visual complexity, with more visual clutter than UK roads. This complexity leads to a broader search pattern in drivers from both nationalities when navigating Malaysian roads [148]. According to Ozkan et al. [173], in a hazard perception test, UK drivers reacted the fastest to hazards, while Chinese drivers were the slowest. Chinese drivers consistently rated all clips as less hazardous than other groups, suggesting that Chinese drivers may have a higher threshold for recognising hazards, possibly due to desensitisation from frequent exposure to high-collision environments [148]. Dey et al. [26] and Goodge et al. [46] highlighted the scarcity of research on culturally aware interfaces in the development of automated vehicles, pointing that AVs should adapt to different cultural preconditions or be clearly designed for specific cultures to enhance their effectiveness. The influence of culture and driving environments extend beyond conventional driving behaviour to the realm of Level 3 cars, though few studies have explored cultural differences in this context. Hergeth et al. [56] studied the effects of TORs on drivers'

trust in highly automated driving among Chinese and German drivers, while Chinese drivers reported significantly higher mistrust in automated driving, the trust ratings did not significantly predict take-over times between the two nationalities. Strle et al. [144] found that US drivers responded significantly slower to TORs than Slovenian drivers, possibly due to higher trust in automation or differences in driving behaviour. The existing research, however, used a one-size-fits-all TOR, ignoring cultural differences that may influence TOR design and thus affect control transition performance. These studies have cemented the need for a more nuanced design of TORs tailored for drivers from different cultures. Failing to do so, Level 3 cars may be designed following a 'one size fits all' approach, ignoring the cultural norms and requirements of specific subgroups.

Driving culture in Israel and the UK

The UK and Israel have distinct driving cultures and regulatory environments [148], which provide a valuable contrast for examining the impact of cultural factors on TOR design. The UK represents a context where drivers are generally accustomed to strict compliance with traffic laws and structured driving behaviour [145]. In contrast, Israel's driving culture is characterised by more aggressive driving behaviour [135], culturally determined informal traffic rules [135] and a different approach to hazard perception [148] than the UK. These differences translate to a higher rate of road incidents in Israel than in the UK. In 2018, Israel reported one fatality from road crashes for every 23,440 residents, compared to one fatality per 36,842 residents in the UK [148]. Research on control transitions in non-Western contexts is scarce, leaving a significant gap in understanding how cultural differences and driving norms shape TOR design. The difficulty in generalising traffic research from Western to non-Western countries [148] underscores the need to explore these areas further. By comparing two distinct driving cultures, with the UK representing high compliance with traffic laws and Israel representing low compliance with traffic laws, we develop TOR designs that are effective across a wide spectrum of driving behaviours.

5.2.4 Summary and research gap

Cultural and age differences, as well as Non-Driving-Related Task (NDRT) types are crucial factors that shape drivers' responses to Take-Over Requests (TORs) in Level 3 vehicles. Research has shown that cultural norms significantly influence driving behaviour and responses to alerts [56, 144, 148, 46], while older drivers face additional challenges during take-over scenarios due to cognitive decline and slower reaction times [117, 93]. Furthermore, NDRT type, particularly visually demanding tasks, can exacerbate difficulties in taking back control [168]. Research has established, therefore, that it is imperative to optimise TOR design to mitigate the challenges drivers face during control transitions. Yet to date, significant research gaps remain unaddressed: **(1)** older drivers, despite being particularly vulnerable during control transitions, are under-represented in studies on TOR design, **(2)** there is no research exploring how cultural factors specifically influ-

ence TOR design, and **(3)** current TOR design research has largely focused on limited NDRTs, overlooking the varying sensory and cognitive demands that different NDRTs impose on drivers in Level 3 cars.

This study is unique as it explores the interplay between these factors, offering insights into how TORs can be better tailored to ensure safe and efficient control transitions across a wide range of driver profiles. By integrating perspectives from different age groups and cultural contexts, our work provides a foundation for developing more adaptive and inclusive TOR designs that cater for the specific needs of diverse driver populations.

We answer the following research questions:
RQ1. Are there age-related differences in the perception of TORs? How are these differences reflected in TOR design requirements?
RQ2. Are there culture-related differences in the perception of TORs? How are these differences reflected in TOR design requirements?

5.3 Methodology

5.3.1 Study design

In each country, three groups of older drivers and three groups of younger drivers were tasked with designing Take-Over Requests (TORs) around four different Non-Driving-Related Tasks (NDRTs): (1) Reading, (2) Watching a video, (3) Chatting with a passenger and (4) Texting. These NDRTs were selected based on their occurrence in previous literature and because they involve varying levels of cognitive and physical demand [108, 72, 141]. Each group comprised of one HCI (Human-Computer Interaction) researcher, who served as the moderator of the session, and two drivers (belonging to the same country and age). See Figure 5.1(B). The study took place in an outdoor car park, with participants seated inside a stationary car. Provided with various materials (e.g., books, smartphones), participants engaged with each NDRT, completed a NASA-TLX questionnaire to assess workload of that NDRT then designed a TOR using the car's interior, windscreen (resembling Head-Up Display. See Figure 3.1), and their chosen NDRT device (e.g., paperback book, smartphone). Each session lasted around 95 minutes. This study was registered and received approval from Glasgow University Ethics Committee (approval ID: 300220195). It took place between June 2024 and August 2024. All participants provided informed consent after receiving a comprehensive explanation of the procedures involved in the study. After transcription, all recordings of the sessions were destroyed and the artefacts kept in a secure storage at the University. We collected the following data:

- **NASA-Task Load Index [55] (NASA-TLX) questionnaire**: completed after engagement with each NDRT, this questionnaire was used to assess the perceived workload of the NDRT (for instance for the Texting NDRT, participants texted for

3 minutes then completed NASA-TLX to assess the workload of Texting). It consists of six items with a 0 to 10 Likert scale (“0" represents low and ”10" represents high). The six items of mental demand, physical demand, temporal demand, effort, performance, and frustration provide a comprehensive measure of perceived workload [54]. The questionnaire was given to the participants in their native language.

- **Qualitative data:** after designing each TOR, participants were recorded describing their designs and the rationale behind them, discussing key features and choices. The pairs within the group also exchanged feedback on each other’s designs. We collected their sketches and designs, documenting them with photos. The collected artifacts and recordings were used in the thematic analysis to gain deeper insights into participant preferences and design decisions.

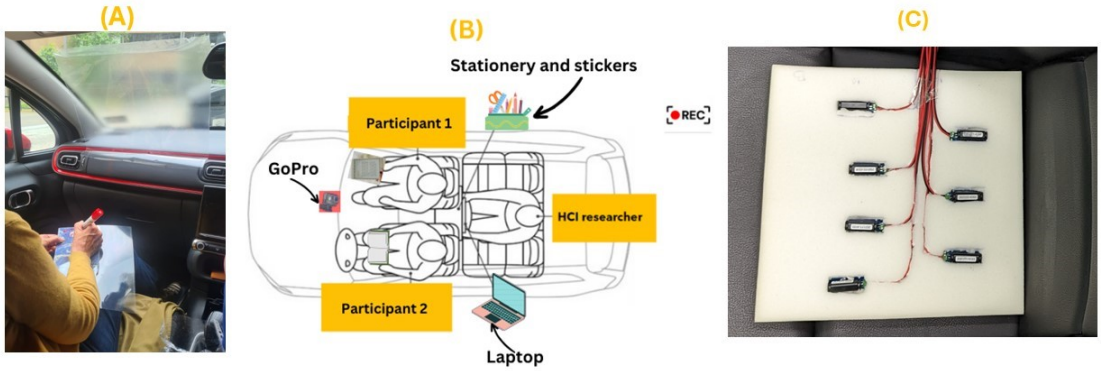


Figure 5.1: (A) A participant designing the TOR on their personal device and HUD laminate. (B) Design study setup with a design group inside a parked car. (C) Foam sheet with 7 actuators placed on the car seat (prior to covering with fabric)

5.3.2 Apparatus

We used a Citroen C3 2019 in the UK and a Seat Ibiza 2018 in Israel. These vehicles were parked in outdoor car parks with dynamic traffic to provide a realistic environment that mimicked actual driving conditions. Participants were provided with laminate sheets as canvases for sketching and visualizing their designs. One sheet simulated the Head-Up Display (HUD), while others could be placed on NDRT devices or elsewhere in the car interior. To enhance the design process, they were provided with sticky notes, markers, and adhesives, enabling them to annotate and adjust their designs interactively. For the tactile component, the study utilised seven actuators from the Actronika HSD MK-I prototyping KIT 4[1]. These actuators were embedded in a foam sheet, which was placed on the car seat and covered with fabric (see Figure5.1 (C)). The actuators provided vibrotactile feedback, set at a frequency of 250 Hz [140], to simulate seat vibrations as part of the

[1]https://www.actronika.com/discover-haptic-technology

Take-Over Request. Participants were encouraged to bring their own devices to perform the non-driving related tasks, such as a Kindle, paperback book, smartphone, or iPad, for the Reading and Watching a Video tasks, allowing for a more personalised and realistic experience. If participants did not bring their own materials, the experimenter supplied necessary items like a book or smartphone for reading and a smartphone or iPad for video watching. For the Texting task, smartphones with pre-loaded Facebook Messenger accounts were provided by the experimenter. For Chatting task, participants chatted with each other and with the moderator. The design sessions were audio-recorded using a Go-Pro camera positioned at the car's tableau to capture the participants' discussions and ideas throughout the process. These recordings were essential for the thematic analysis that followed, helping researchers understand the decision-making and design considerations of each group. Photographs of the participants' designs were taken by the moderator. These images documented the visual and physical aspects of the designs, providing a reference for further analysis and comparison across different groups and tasks.

5.3.3 Participants

We recruited 24 older (60+ years old) and younger (22-32 years old) drivers from the UK and Israel , with 12 from each country: the UK (younger drivers (n=6; Mean age=25.3yrs, SD=1.6yrs; Males=4, Females=2), older drivers (n=6; Mean age=68.3yrs, SD=4.1yrs; Males=2, Females=4)) and Israel (younger drivers (n=6; Mean age=28.5yrs, SD=2.1yrs; Males=3, Females=3), older drivers (n=6; Mean age=63.3yrs, SD=4.6yrs; Males=4, Females=2)). All participants were active drivers and had held a valid driving license for at least the past three years. Older drivers from both countries reported normal to corrected-to-normal vision and hearing and had no compromised sense of touch. All participants declared they knew about the concept of automated cars but had never experienced them. We recruited participants via internal university e-mail lists and through personal contacts and social media.

5.3.4 Procedure

Upon arrival participants signed consent and were seated in the front seats of the car. Next, they were provided with a comprehensive overview on Level 3 car technology with focus on control transitions. Participants were instructed to engage in a Non-Driving Related Task while imagining that the vehicle is operating in automated mode. Following each engagement, they were told that there is a non-critical road hazard and the time budget[2] of the TOR was 15-20 seconds. Subsequently, participants were tasked with designing a take over request to effectively communicate the transition of control from the automated system back to the driver. They were advised that TORs could include

[2]TOR time budget was defined during which the participant should identify the hazard and regain manual control to avoid a potential crash. In the literature, the duration of the TOR budget ranges between 5 seconds (critical) [71] to 20 seconds (non-critical) [93, 71]

visual, auditory, or vibrotactile feedback, and were encouraged to freely and creatively integrate these or other components. Design sessions ran for $\approx$ 95 minutes each ($\approx$ 40 contact hours in total), split into: (1) **NDRT engagement:** 3 minutes of NDRT engagement (i.e., reading, watching a video, chatting or texting), (2) **Questionnaire:** the NASA-TLX questionnaire was administered to assess perceived workload of the NDRT [55], (3) **Ideation:** 10 minutes of TOR ideation. Participants independently sketched their designs and placed them inside the car and on the device laminate to visualise the TOR features. To demonstrate seat vibrations, the moderator activated actuators to simulate continuous or intermittent patterns. Participants could then suggest vibration patterns they found most effective, (4) **Design presentation:** each participant described their designs aloud and gave feedback on the other participant's designs for 8 minutes and (5) **Documentation:** 2-3 minutes of photographing an collecting the designs and preparing the setting for the next NDRT. This process was repeated for each of the four NDRTs, allowing participants to consider each in detail. Groups engaged with NDRTs in a counterbalanced order and participants switched seats after the second NDRT to give an equal opportunity for them to be in the driver's seat. The experimenter was fluent in English,Arabic and Hebrew, eliminating the need for translation and allowing participants to express themselves in their native language. Participants received £10 Amazon voucher at the end of the experiment as a compensation for their participation. The collected designs formed the basis for a taxonomy of TOR features considering the different NDRTs, which were used to synthesise versatile TORs by identifying overlapping features between designs.

5.4 Results

5.4.1 NASA-TLX overall workload

To gain knowledge of how demanding each NDRT type might be, participants completed a TLX assessment after engaging with each one. The results were analysed using R Studio 2022.07.2 Build 576 using the lme4 [10], lmerTest [83] and report [97] Rpackages. A linear mixed model was initially fitted to assess the effects of NDRT type, age group (younger vs. older), and country (UK vs. IL), including their interactions, on the overall NASA-TLX score. The interaction terms were not significant, so backward elimination was performed, resulting in a final model with three fixed effects: NDRT, age group, and country. Random intercepts for participants were included. The analysis revealed a significant main effect of NDRT *($F(3,68.85)=18.51$, $p < 0.0001$, $\eta^2 = 0.45$)* and age group *($F(1,21.5)=6.9$, $p = 0.015$, $\eta^2 = 0.24$)* while the effect of country was not significant *($F(1, 21.495) = 1.240$, $p = 0.278$)*. Younger drivers had significantly lower NASA-TLX scores than older drivers (Estimated mean difference = -0.907). *Post hoc* analyses with a Bonferroni adjustment confirmed differences between NDRTs. The Video task had significantly lower NASA-TLX scores compared to Chatting ($p < 0.0001$), reading ($p = 0.0015$), and Texting ($p < 0.0001$). The Reading task showed significantly lower scores than Texting

($p = 0.0194$). However, no significant differences were found between Chatting and Reading or Chatting and Texting. Results are presented in Figure 5.2.

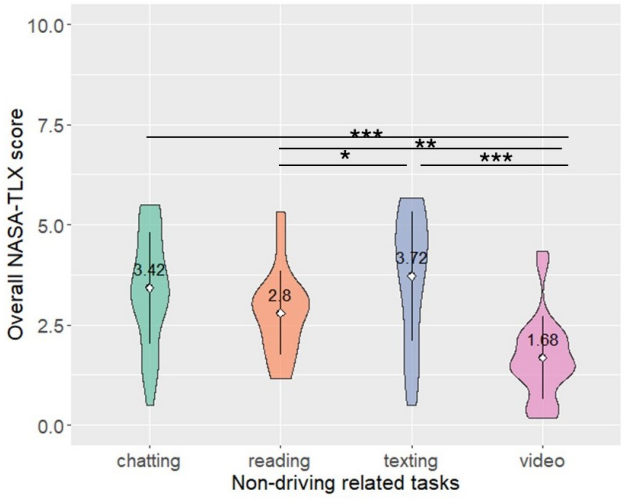


Figure 5.2: Overall NASA-TLX workload score for each of
the NDRTs. The dots represent mean values.
*** - P<0.001,
** - P<0.01, *- P<0.05.



Figure 5.3: Red light band across the dashboard for the visual component of the take over command.



Figure 5.4: Risk meter showing hazard criticality, integrated into the HUD takeover command, based on Israeli participants' requirements.

5.4.2 Theme Extraction

Recordings of participant discussions were transcribed using Sonix.ai [3] and one author reviewed the transcripts against the recordings to correct any errors. The data were then imported into NVivo[4] for an inductive, data-driven, thematic analysis [17]. Initially, one researcher identified distinct codes within the data. Two researchers then classified these codes into broad thematic categories based on their similarities. Any disagreements about code placement were discussed and resolved until consensus was reached. The same researchers developed themes by combining similar codes into broader themes. If significant overlap was found between themes, the researchers considered merging them. Throughout this process, disagreements were resolved through discussion, ensuring that the final themes accurately reflected the data. This process was done for each group separately (UK older drivers, IL older drivers, UK younger drivers and IL younger drivers). Overall, seven themes were extracted, four of these (themes 1-4) were common to all study groups and three emerged only for specific groups (themes 5-7). This section presents themes identified across all groups and features participant designs that illustrate them.

[3]https://sonix.ai/
[4]https://help-nv.qsrinternational.com/14/win/Content/welcome.htm

Theme 1: Mental Resource Allocation and NDRT Engagement

Participants (N=24) reflected on how they allocated cognitive, sensory, and physical resources while engaging with different NDRTs. Discussions highlighted varying levels of mental demand, visual attention, and physical involvement, depending on the nature of the NDRT.

Chatting with a passenger. This was likened to watching a video in terms of observing facial expressions, and described as cognitively similar to texting because it involved processing the other person's words and formulating a response. Par21 (IL,28yrs) elaborated *"all my senses were engaged [..] cognitively I could not think of anything other than our conversation"*. Participants preferred to look at the person they were conversing with. Par1 (IL,68yrs) explained *"you are engaged with facial signals and information from the person you're speaking with, and you're also getting ready for the next thing you want to say"*. The physical engagement while chatting was evident as all participants shifted their bodies away from the driving direction to face the passenger they chatted with. Par5 (UK,68yrs) stated *"the big issue is body position. It is risky. If we're in a deep discussion, I'm gonna have eye contact with you and shift body posture towards your direction to express and observe body language"*. Participant reflections on this NDRT highlight their physical and cognitive shift away from the driving task, potentially compromising situational awareness and the ability to respond quickly to road hazards.

Texting. Given its manual-visual nature, most participants agreed that texting was more resource-intensive than reading and watching a video. Par2 (IL,74yrs) noted that the interactivity in texting made it more difficult to relinquish the non-driving task, *"you might be more tempted to keep your attention on it and want to respond quickly. So it's different in that way, because a book or a video you can just forget about and put down"*. Par23 (IL,26yrs) added *"the visual search for letters and coordination of typing and reading texts burdens the visual resource"*.

Reading. Par1 (UK,68yrs) described their engagement with reading using an e-reader (e.g. Amazon Kindle) as *"probably I'd want as much stimulation to get me out of, I could be quite deeply absorbed in that [reading]"*. Reading was considered to occupy working memory, as participants mentioned the need to remember the last sentence they read to revisit it after a take-over manoeuvre was complete. Par17 (IL,60yrs) said *"I would want to remember where I stopped reading to be able to revisit the same sentence. This is mentally demanding"*.

Watching a Video. Most participants agreed that it was the least demanding of the NDRTs. Par3 (UK,63yrs) elaborated *"I'm lightly engaged with it, it's more passive. I can stop that at any point"*. Par24 (IL,30yrs) reiterated *"the movie is superficial. Your attention is not captured in the same way as texting for example"*.

Mapping to RQ1 and RQ2: Across both cultures, both age groups reported that the NDRTs varied in their cognitive, sensory and physical demand.

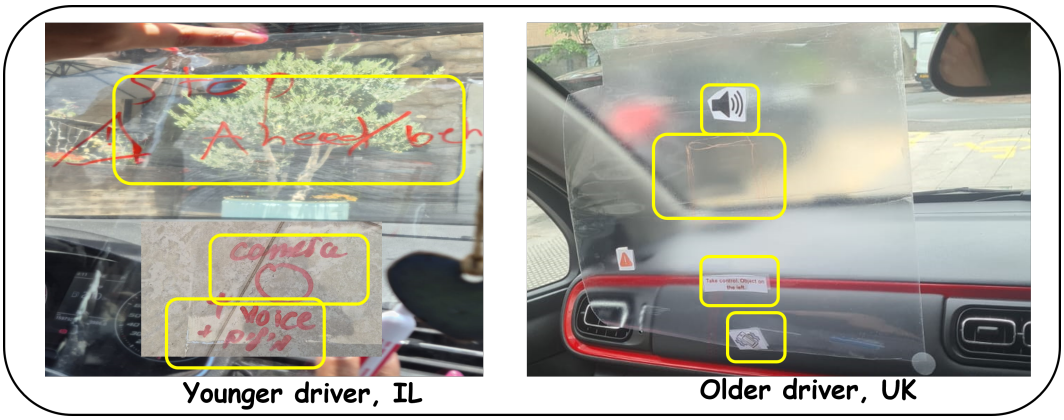


Figure 5.5: Proposed designs by younger (left) and older (right) drivers from IL and UK respectively. Laminate on windscreen represents Head-Up-Display. Both the type and location of the hazard were conveyed through text. For example: "Take control. Object on the left" (Older driver, UK). Hazard location was indicated on the HUD with a flashing frame, dynamically adjusting as the vehicle approached.

Theme 2: TOR visual component placement and content

Participants placed significant emphasis on the importance of the TOR's visual component, focusing on where messages should be displayed, the type of information conveyed, and the appropriate tone for these communications.

Essential messages to communicate. The majority of participants (N=20) preferred a verbal or written message over an icon for the TOR. Twenty two participants stressed the importance of the TOR being both informative, providing clear take over command and the hazard type, and positional, indicating the hazard's location. For instance, Par21 (IL,28yrs) suggested, "I would want to see *"Take over! Object on the left [pointing at the windscreen to indicate location of the message]"*. In terms of positional information, Par24 (IL,30yrs) noted, *"If the system can tell me where to look, this could save me effort"*. Participants also recommended complementing the written message by a frame on the windscreen (e.g., HUD- Head Up Display). Par18 (IL,60yrs) suggested, *"A red frame around the hazard would be helpful to attract my attention"*. See Figure 5.5 for sample designs. However, when proposed, opinions varied regarding a dynamic countdown bar indicating time left for a safe takeover. Older drivers were more divided; Par19 (IL,72yrs) felt, *"It might make people think I've got time, as opposed to responding immediately"*, whereas Par2 (UK,74yrs) explained, *"I do not mind having it as I need time to put my phone away, so I need to know how much time is left"*. In contrast, the majority of younger drivers (N=11) favoured the countdown bar, as it helped them manage the stress of the takeover process.

Tone of the verbal message. The majority of older drivers (N=9) advocated for an authoritative verbal TOR message. Par3 (UK,63yrs) explained *"I don't want to hear "please take over control". It's too soft"*, Par20 (IL,74yrs) added *"I want it minimalist and authoritative. No fuss, it's a safety matter"*. Younger drivers however had a more even split in their requirements. Six of the younger drivers preferred a more polite TOR while others preferred an

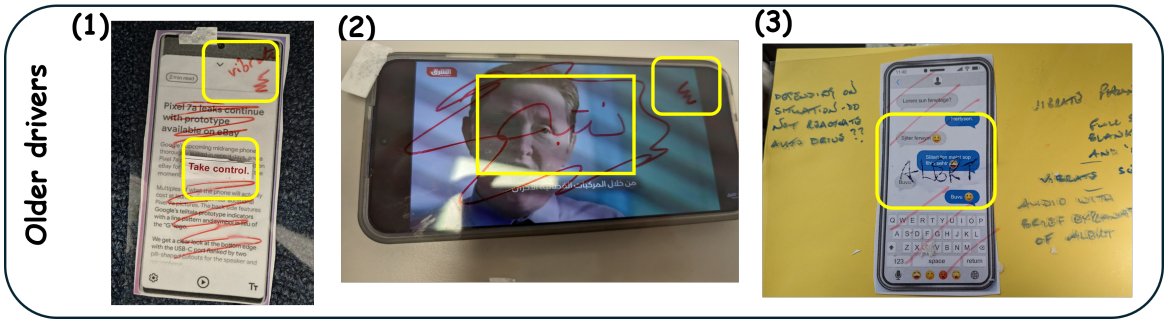


Figure 5.6: For NDRTs involving handheld devices (Reading (1), Watching a video (2), and Texting (3)), older participants consistently designed a unified interface. The "take over" command was prominently displayed at the centre of the device screen, immediately disabling the NDRT and obscuring the screen with text (e.g., "take over" or "alert"), accompanied by device vibration.

authoritative command. Par23 (IL,26yrs) stated *"why not be polite? if the car is becoming more personalised I would want it to be polite to me and say "please take over""*.

NDRT suspension. All older drivers (N=12) wanted the TOR to be displayed in the middle of the device on which the NDRT was displayed, suspending the NDRT and obstructing the screen. Par3 (UK,63yrs) described her design approach, stating, *"I was working on my Kindle so I've tried to make the TOR take as much of the face of the Kindle as possible because you're quite deeply involved in reading"*. Par16 (IL,62yrs) emphasised this by saying, *"I want the screen to go blank with only the TOR displayed in the middle"*. This ensures that the driver's full attention is redirected to the driving task. See Figure 5.6 for sample designs. Younger participants, on the other hand, had more variance and raised the theme of NDRT interruptibility as described in Younger drivers: interruptibility of the NDRT. They advocated for a more flexible TOR, displayed at the periphery of the NDRT- device screen, not obstructing it. Par25 described *"I want it to be at the top like a text-message notification that stays on"*. For the Chatting-with-passenger NDRT, which lacks a handheld device, participants across groups suggested the employment of the car interior. For example, Par5 (UK,68yrs) said *"a red light band across the dashboard for the take over command would draw my attention back to driving"* (see Figure 5.3 for an illustration).
Mapping to RQ1 and RQ2: Across both cultures, age-based differences emerged, as older drivers preferred more obstructive TOR displays while younger drivers sought peripheral placements.

Theme 3: Modalities of the TOR

All older drivers preferred a tri-modal take over request that included visual, auditory and vibrotactile components. Par20 (IL,74yrs) stated *"I would want as many senses as possible to be alerted"*. This preference was justified by recognising the early signs of sensory deterioration. Par19 (IL,74yrs) explained *"my hearing and vision are not as sharp anymore. Relying only on one is risky for me"*. The majority of younger drivers (N=10) preferred bi-

modal take over requests that combined a visual alert with either a vibrotactile or auditory component. Par22 (IL,28yrs) explained, *"Having all three modalities will unnecessarily startle me"*. Similarly, Par23 (IL,26yrs) noted, *"I want to reserve the tri-modal TOR for highly critical events"*. Preferences for the second modality varied according to each participant's perceived sensory sensitivity. For instance, Par21 (IL,28yrs) stated, *"Vibration causes me discomfort and stress. I prefer to rely on my hearing"* while Par23 (IL,26yrs) said *"I will always be in contact with the seat while driving, I definitely want it to vibrate"*.

Mapping to RQ1 and RQ2: Across both cultures, evident age-related differences in sensory preferences for TORs emerged. Older drivers preferred a tri-modal approach, while younger drivers opted for bi-modal TORs.

Theme 4: Personalising and anthropomorphism

Participants expressed a clear preference for avoiding avatars or facial features for the visual component of the TOR, as these could be distracting. Par7 (UK,27yrs) *"I do not see the benefit of having an avatar. This gives the impression that the TOR is not serious"*. Instead, they emphasised the need for straightforward instructions without unnecessary elements. Par4 (UK,72yrs) said *"No fussiness, you know, to distract from what the major problem is"*. Participants also favoured having their first names as part of the TOR. Par3 (UK,63yrs) noted the effectiveness of using names in alerts, stating, *"If it was to be humanised at all, it might use your name because your name pulls you back quite quickly. For example: "[Driver's name]! Take control. Work zone ahead"*.

Mapping to RQ1 and RQ2: Across both cultures, both age groups avoided avatars or facial features for the visual component of the TOR.

Theme 5 (specific to older drivers from UK and IL): Ageing, driving safety and legality

Older participants across both cultures reflected on how ageing affects driving safety. One participant highlighted the impact of ageing on their driving confidence: *"I feel that my driving is more cautious and I become more panicked quite quickly"* Par3 (UK,63yrs). Additionally, the effects of ageing on cognitive and physical abilities were discussed, Par13 (IL,62yrs) shared *"vision and hearing loss become imminent at my age and make reaction times slower"*, Par2 (UK,74yrs) added *"Your powers definitely are waning, the neural pathways are maybe a bit more clogged affecting decision making"*. Catering for older drivers' deterioration in abilities in the design of TORs would help mitigate the take over challenge for this group. Referring back to Theme 3, older participants' preference for tri-modal TORs likely stems from their acknowledgment of sensory decline.

Legality. The discussion also touched on the upcoming changes in legality with Level 3 autonomous vehicles, where activities like texting and shifting body posture away from driving would become legal. Par15 (IL,61yrs) raised the concern *"I would be reluctant to drive knowing drivers can text. I think it's too immersive and should be restricted"*. These changes necessitate efficient Take Over Requests to ensure safety. Par18 (IL,60yrs) em-

phasised, *"If these activities are legal, it's the manufacturer's responsibility to design fail-safe TORs"*. Highly engaging tasks like texting and chatting with passengers, which significantly divert attention, should be carefully considered.

Versatility. Older participants across the two countries agreed that the windscreen-TOR (i.e. HUD) and device-TOR designs should be consistent across all NDRTs involving a device or a defined interface (e.g., texting, reading from Kindle, watching a video). Par5 (UK,68yrs) explained *"The TOR message should have the same colour, location and content across all devices. Otherwise I would be confused"*. This uniformity would simplify processing the information without needing to adjust for each task type. Par4 (UK,72yrs) iterated *"a different TOR for each task is like clutter which you need to filter, whereas if it's just specific and to the point and it's the same each time, you're recognising it"*, Par3 (UK,63yrs) added, *"I don't want it to vary, so I get used to it and not need to interpret each time"*.

Mapping to RQ1: Age-related sensory and cognitive decline influenced older drivers' preferences for tri-modal TORs and consistency of TORs across non-driving related tasks.

Theme 6 (specific to Israeli participants): Driving culture and compliance with traffic laws

Participants (N=12) described driving in Israel to be reckless and drivers often exhibit a lack of adherence to traffic laws. Par21 (IL,28yrs) mentioned *"the TOR system will probably work more in IL than in western countries because the driving culture leads to more hazard occurrences"*, Par22 (IL,28yrs) elaborated *"if the system issues a danger signal which is according to the mentality of the driver not dangerous, they will most likely ignore it"*. Participants suggested law enforcement to address these issues. Par16 (IL,62yrs) stated *"response times to the take over requests and quality of take over should be recorded by the system, if below a certain threshold one should be penalised or undertake training"*.

Sensitivity to alerts. The ongoing regional conflict in Israel has led to a unique sensitivity to alerts, with people often being hyper- or hypo-sensitive to alerts. Par18 (IL,60yrs) highlighted *"we live in a reality of danger alerts, TORs must balance urgency with credibility"*. Participants (N=9) suggested designing TORs with context-specific sensitivity, Par17 (IL,60yrs) said *"it would be helpful to have the level of urgency of the hazard displayed. It will help manage stress levels"*. *"I want to know how hazardous is the hazard, if it is a small tree branch I should not panic"*- Par23 (IL,26yrs). See Figure 5.7 for a sample design.



Figure 5.7: Younger driver's (IL) design showing a meter for hazard criticality.

Mapping to RQ2: Cultural differences were prominent, with Israeli participants emphasizing the impact of local driving habits and sensitivity to alerts on TOR design, suggesting a need for real-time risk assessment.

Theme 7 (specific to younger drivers from the UK and Israel): Interruption management

Nine out of 12 younger drivers preferred having control over the interruptibility of non-driving related tasks after a non-critical takeover request, wanting the option to suspend the task at their own convenience. The rest (N=3) wanted the task to be immediately blocked following the takeover request, similar to the preference of older drivers. Par21 (IL,28yrs) said *"if it is legal to use my smartphone in Level 3, I would want to be trusted to stop the task after the TOR by myself and not have this done for me"*. Par8 (UK,27yrs) stated *"if it will take me one more second to finish the text that I was typing, I would want to finish it and then take over"*. See Figure 5.8 for sample designs. To manage the interruptibility of the task, participants favoured having a countdown timer for the TOR, Par10 (UK,23yrs) elaborated *"a countdown timer will help stay in proportion and know what decision to make"* (see Figure 5.8 (2)). They opted for a TOR that is gradual rather than abrupt. Par25 (IL,32yrs) elaborated *"..if for some reason I did not take control and the time is elapsing I would want a fiercer alert"*.

Mapping to RQ1: Younger drivers favored control over NDRT interruption, preferring gradual alerts, while older drivers preferred immediate interruption, reflecting an age-based difference.

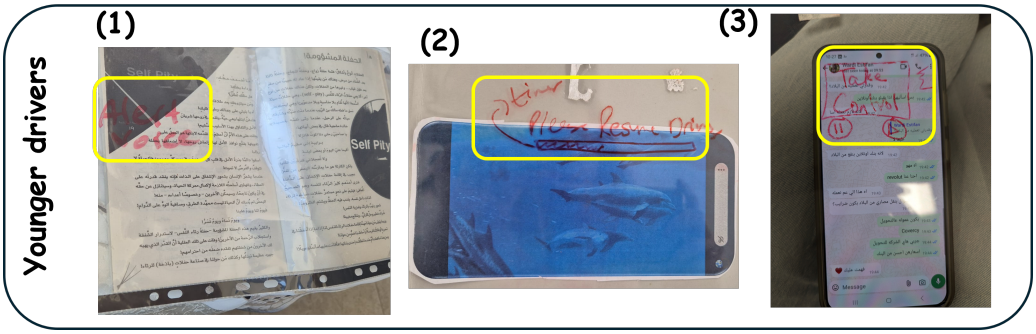


Figure 5.8: For NDRTs involving handheld devices (Reading (1), Watching a video (2), and Texting (3)), younger drivers preferred a "take over" command prominently displayed in red either at the top of the device's screen or in the periphery.

5.4.3 Taxonomy development

A taxonomy provides a visual framework to understand the relationships and hierarchy of key TOR features, such as content, modalities, and placement, helping optimise design decisions to meet diverse user needs. To develop a taxonomy of key TOR features across the four NDRTs, we conducted a theoretical thematic analysis [17] using NVivo to identify features such as TOR content, modalities, or placements discussed during design sessions. Using the approaches described by Al-Taie et al. [3], we developed a hierarchical taxonomy. One researcher initially coded these features from the designs, focusing exclusively on HCI (Human Computer Interface) aspects, while another researcher grouped

the codes into themes such as modality or placement, refining them iteratively. Four key themes representing HCI features emerged: Communicated message, Modality, Specific placement and Technology. To clarify feature relationships (e.g., best placements for specific modalities), we established a taxonomy based on the order participants typically followed in their design discussions. Codes were relabelled according to higher-level themes, and two researchers independently assessed their frequency in sequences, confirming no discrepancies. The final taxonomy layers, populated with features and their frequency appearance in the designs, enabled us to identify commonalities and differences across TOR designs for various NDRT types. **Our final taxonomy consisted of five layers: (1) NDRT type, (2) Communicated message, (3) Modality, (4) Specific placement and (5) Technology.**

We developed two taxonomies: one for older (UK and IL) and one for younger drivers (UK and IL). Each was split into two: one for handheld device-NDRTs and another for Chatting with a passenger NDRT, resulting in four taxonomies. We selected this division because it highlights the areas where the most discernible differences were identified. Figures 5.9 and 5.10 present tree diagrams outlining the taxonomy of interface components. This visualisation demonstrates how the handheld devices were incorporated as part of TOR interfaces, which can assist designers in selecting suitable interface modalities and messages. The taxonomy also addresses RQ1 by illustrating how interface designs vary between the two age groups. For example, while all older drivers consistently preferred tri-modal TORs to communicate the takeover command, younger drivers leaned towards bi-modal TORs, with all preferring a visual component, 50% choosing vibrotactile, and 67% auditory (see "Modality" layer in Figures 5.9 and 5.10). Additionally, younger drivers' preference for including a countdown timer is shown in an added branch under the "Communicated message" layer in Figure 5.10.

5.4.4 Synthesis of design concepts

We present five versatile TOR design concepts, each representing key findings from our design sessions. These concepts (see Table 5.1) serve as case studies to discuss our results. The concepts were synthesised by using the taxonomy to identify primary TOR

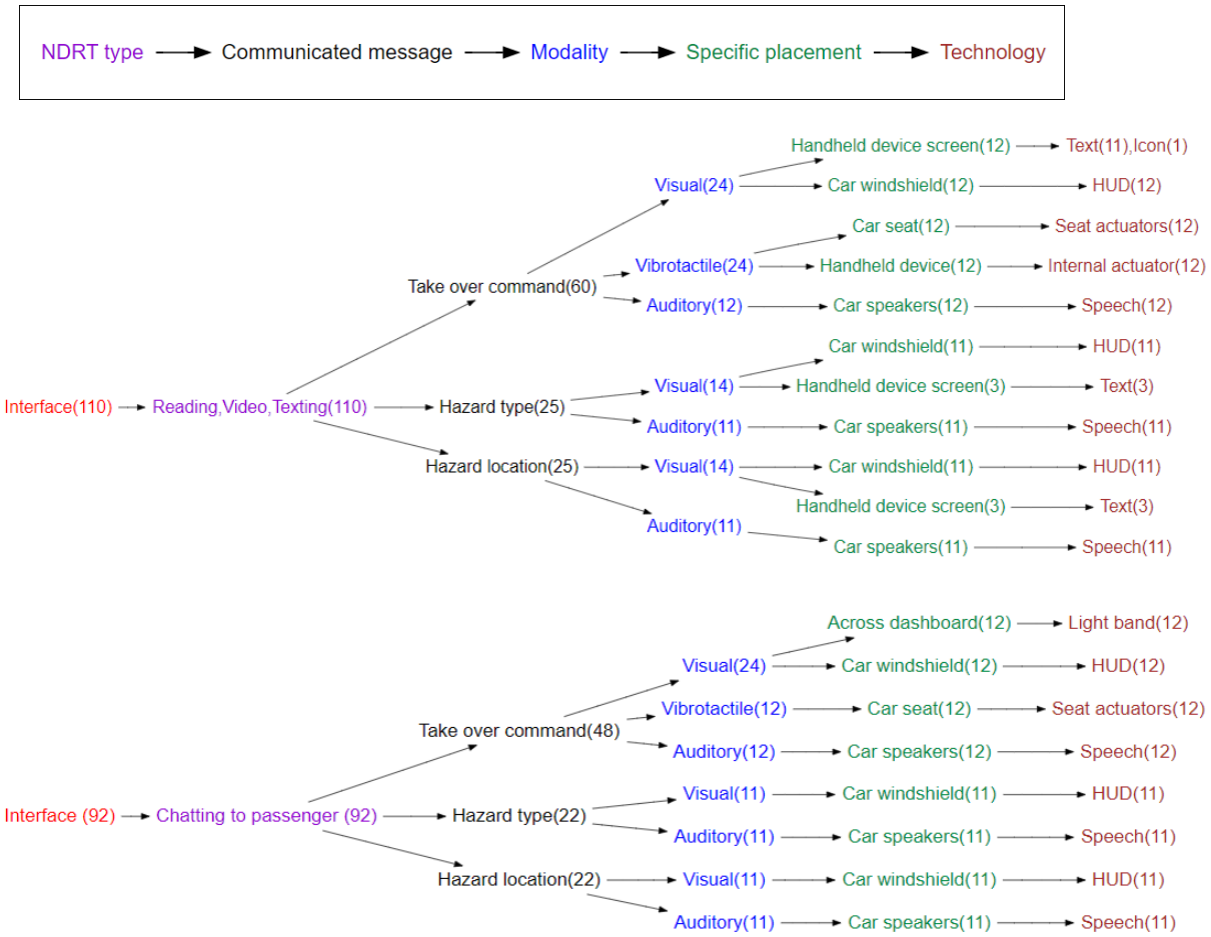


Figure 5.9: Older drivers' taxonomy of TOR features for all the NDRTs including a device (upper tree: Reading, Watching a video and Texting) and Chatting with a passenger (bottom tree). Each branch is labelled with: TOR feature (frequency of appearance).

components that overlapped across NDRTs and participant groups. We reviewed participant designs and emergent themes to pinpoint secondary components that complemented the primary ones, refining the designs iteratively until achieving complete TOR concepts. Table 5.1, Figures 5.11 and 5.12 provide a detailed description of the rationale behind each TOR concept, outlining its functionality, the data sources from which it was derived, and its key features. The concepts in Table 5.1 are: (1) **Context-Aware TOR** enhances situational awareness by displaying take over command, hazard type and location dynamically on the HUD. This is informed by Theme 2 and taxonomy layers related to "Communicated message" and "Technology" from taxonomy 5.9 and 5.10. (2) **Dual TOR** synchronizes alerts across NDRT devices and the in-vehicle information system, incorporating engagement-detection technologies such as eye tracking. This is guided by Themes 2 and 7, and the "Specific placement" layers of taxonomy 5.9 and 5.10. (3) For older drivers, the **Disruptive TOR** employs tri-modal alerts and suspends NDRTs to ensure an immediate response, reflecting insights from Theme 2 and taxonomy 5.9. (4) Tailored for younger drivers, the **Gradual TOR** introduces a 3-phased TOR with countdown timers and customizable cues. This is shaped by Theme 7 and taxonomy 5.10.

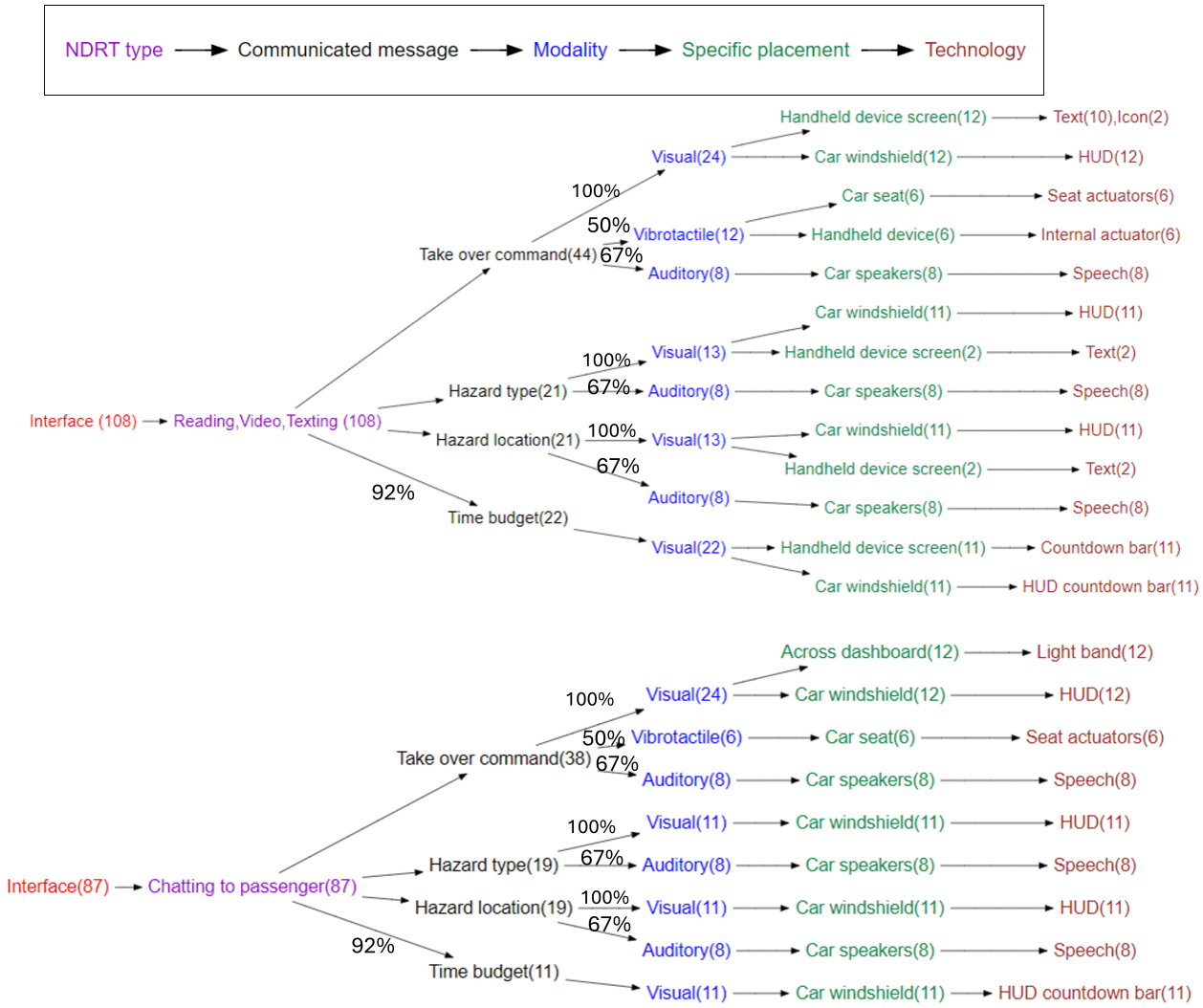


Figure 5.10: Younger drivers' taxonomy of TOR features for all the NDRTs including a device (upper tree: Reading, Watching a video and Texting) and Chatting with a passenger (bottom tree). Each branch is labelled with: TOR feature (frequency of appearance). Percentages on the branches indicate the proportion of participants wanting that feature. For example: 92% of the younger participants wanted to include a time budget.

(5) The **Risk-Aware TOR** features a dynamic Risk-Meter that provides real-time hazard assessments, addressing cultural differences in risk perception as highlighted in Theme 6 and feedback from Israeli participants. Figures 5.11 and 5.12 illustrate the concepts as designs on the NDRT-device and HUD. The HUD displays translucent information, ensuring road visibility is maintained.

5.5 Discussion

Our work lays the groundwork for designing culturally- and age-sensitive control transition TORs, significantly expanding previous research. New findings from older and younger drivers in two cultural contexts— the UK (high traffic law compliance) and Is-

TOR Design Concept	Description	Source (Theme,Taxonomy,Design)	Key Features
Context-Aware TOR	Detailed TOR providing a takeover command (displayed on HUD and NDRT-device), hazard type (displayed on HUD), and location (displayed on HUD). The hazard is highlighted with a flashing frame that adjusts as the driver approaches.	Theme 2, Taxonomies 5.9 and 5.10 "Communicated Messages" and "Technology" layers, Designs: Figure 5.5	Enhances situational awareness, allows drivers to assess risks better, and make informed decisions.
Dual TOR	Repurposing personal devices to issue TORs alongside in-vehicle information system. The system can detect the driver's engagement with the NDRT on the personal device, for example through eye tracking metrics (e.g., gaze) or touch sensors, and automatically trigger TORs that are synchronised across both platforms.	Themes 2 and 7, Taxonomies 5.9 and 5.10 "Specific placement" layer, Designs: Figures 5.6 and 5.8	Provides redundancy and increases response likelihood.
Disruptive TOR (Older Drivers)	A tri-modal, Dual and Context-Aware TOR that prominently displays a flashing red command in the center of the handheld screen, instantly suspending the NDRT and blocking the device's display. The alert is reinforced by simultaneous seat and device vibrations, along with an auditory message. See Figure 5.11.	Theme 2, Taxonomy 5.9, Designs: Figure 5.6	Suspension of the NDRT with multiple sensory inputs (vibration, audio, visual) prompts immediate response from the driver.
Gradual TOR (Younger Drivers)	A 3-phase Context-Aware and Dual TOR. Phase 1, active engagement with NDRT. Phase 2, TOR is issued alongside a countdown timer (on HUD and NDRT-device's screen) indicating the remaining safe takeover time. Feedback through auditory and vibrotactile cues varies based on participant preferences (Figure 5.10). If a takeover action is detected (e.g., using eye-tracking metrics, steering, or braking [71]), the TOR ceases, allowing the driver to resume the NDRT after the manoeuvre. Phase 3, if no action is taken within the specified time, the system issues a Disruptive-TOR, suspending the NDRT. See Figure 5.12.	Theme 7, Taxonomy 5.10, Designs: Figure 5.8	Gradual takeover process with countdown timers and customisable alerts, offering more control over NDRT interruption
Risk-Aware TOR (Israeli Drivers)	Includes a Risk-Meter that displays real-time risk assessments, factoring in driver state, hazard type, and environmental conditions on both the HUD and personal device. See Figure 5.4.	Theme 6, Designs: Figure 5.7	Assists drivers in gauging urgency and managing stress, especially in regions with altered alert sensitivity.

Table 5.1: Summary of synthesised TOR design concepts based on participant designs, thematic analysis and taxonomies.

rael (low traffic law compliance)—, across a range of NDRT types, challenge previous TOR design assumptions.

RQ1: Are there age-related differences in the perception of TORs? How are these differences reflected in TOR design requirements?

Older Drivers. Our study reinforces previous findings on older drivers' preference for tri-modal TORs [64] and their recognition of age-related declines in cognitive and physical abilities [141, 117], while extending these insights to a cross-cultural context. Participants from both the UK and Israel emphasised the need for consistent, straightforward TORs that minimise real-time interpretation. Consistent with Huang et. al [64], all older drivers preferred a tri-modal system with visual, auditory, and vibrotactile components, seeking redundancy to compensate for sensory impairments [77]. While a tri-modal approach seems beneficial for older drivers, it raises concerns about potential information overload [166]. Given that older adults may also experience slower cognitive processing [117], the simultaneous activation of three sensory modalities might overwhelm them, leading to delayed, rather than improved, responses. Therefore, it is crucial to evaluate tri-modal TOR systems and balance the need for multimodal alerts with the risk of

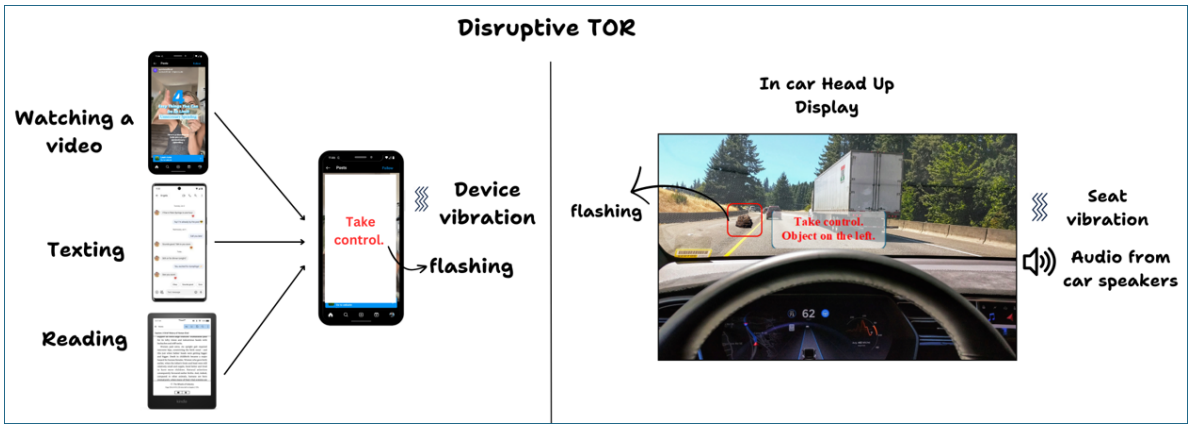


Figure 5.11: "Disruptive TOR" interfaces. Left: the take over command issued on the NDRT device. Right: the HUD design of the TOR.

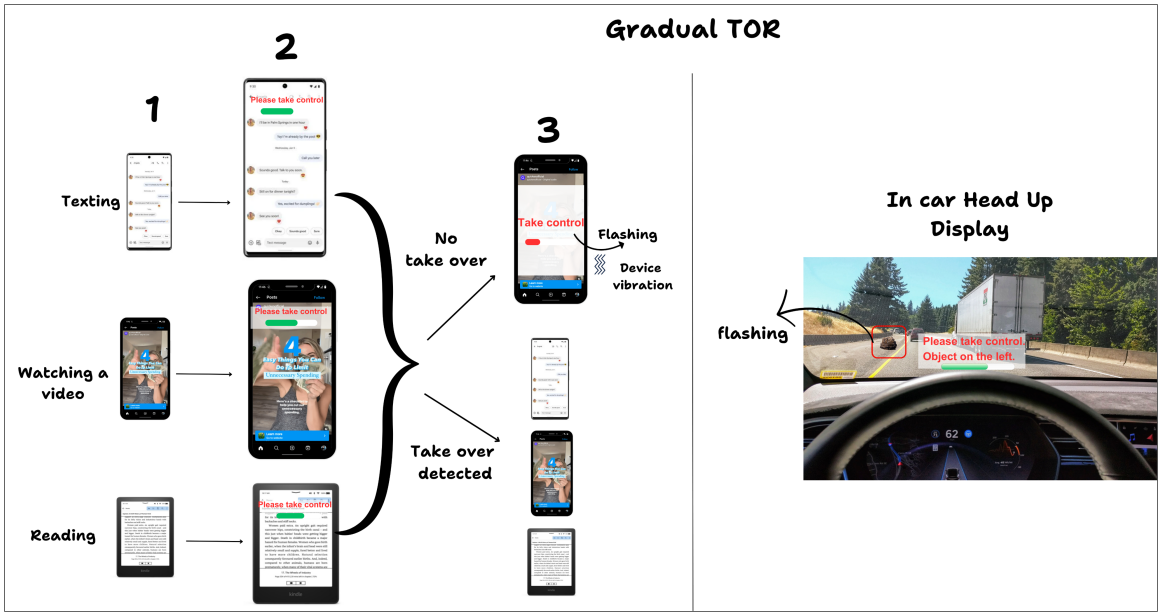


Figure 5.12: "Gradual TOR" interfaces. Left: the take over command issued on the NDRT device. Right: the HUD design of the TOR.

over-stimulation. One approach could be to allow drivers to adjust the intensity of each modality or to prioritise certain modalities based on the individual's sensory strengths and weaknesses. Older drivers preferred **Disruptive-TOR**s (see Figure 5.11 and Table 5.1) that immediately suspended any ongoing NDRTs, reflecting their cautious approach to driving [52, 92, 102] and their prioritisation of safety [141] over task completion. This preference suggests greater risk aversion and awareness of delayed response risks. However, abrupt task suspension, as opposed to gradual, could impose a cognitive cost [160] on the driver impairing their take-over performance. Future research should explore these potential cognitive costs to ensure that the safety benefits of immediate task pausing do not inadvertently impair driver performance.

Younger Drivers. A key finding reveals that most younger drivers preferred managing task interruptions themselves, valuing the ability to pause or continue non-driving tasks at their discretion. This highlights a novel shift towards greater driver autonomy

in Level 3 cars. This could be attributed to younger drivers' fundamental expectations of trust in their ability to manage multiple tasks [18], reflecting a broader trend in which they favour customisation and autonomy over rigid systems [98]. The willingness to prioritise task completion over immediate response may be related to younger drivers' tendency toward risk-seeking behaviours [102], raising important questions about how TOR systems can balance the need for safety with the driver's sense of autonomy. To manage the interruptibility of tasks, participants suggested the introduction of a countdown timer as part of a **Gradual-TOR** (see Figure 5.12 and Table 5.1). A countdown timer has been proposed as part of a TOR in previous literature [57]. While it can help drivers smoothly transition out of non-driving tasks, mitigating abrupt cognitive shifts, it must be carefully designed to ensure that this gradual approach does not delay the necessary return to full driving focus. Wintersberger et. al [157] proposed an approach that could leverage the design of gradual TORs by issuing them at task boundaries. For example, instead of interrupting during typing, the TOR would be triggered after a text message is sent. This approach improves task flow and reduces driver stress but is only feasible when a task boundary can be reached quickly. Most younger drivers preferred bi-modal TORs, combining visual alerts with either vibrotactile or auditory cues, to avoid startling and reserve tri-modal alerts for critical situations. This aligns with studies showing that while multimodal alerts improve response times, they can also increase urgency and annoyance [120, 64]. Preferences for the secondary modality (alongside the visual) also varied among younger drivers based on individual perceived sensory sensitivities. The variation in younger drivers' preferences underscores the need for customisable TOR system, unlike older drivers, who may benefit from a standardised tri-modal TOR.

RQ2: Are there culture-related differences in the perception of TORs? How are these differences reflected in TOR design requirements?

This is the first study to investigate TOR design in the context of driving cultures with varying traffic law compliance. Israeli participants agreed that driving culture in Israel is characterised by lower compliance with traffic laws and a tendency toward more reckless driving, aligning with the findings of Lipaz-Shamoa Nir et al. [135] and Vera Skvirsky et al. [139]. These factors could result in (1) drivers circumventing TORs and (2) more frequent TOR activation due to increased road hazards. Participants, therefore, suggested that TOR systems in Israel may need to adopt a more assertive approach, potentially integrating mechanisms to monitor driver compliance with TORs and incorporating penalties for delayed or insufficient responses. Research by Sezgin et al. [134] supports the idea that integrating monitoring and feedback systems can reduce risky driving behaviours. However, this raises ethical questions about the role of autonomous vehicles in enforcing behaviour, as discussed by Goodall et al. [45] and Goodison et al. [48], who warn of potential liability and autonomy issues. The ongoing regional conflict in Israel, coupled with the frequent danger alerts resulting from this, has caused drivers to become either hyper- or hypo-sensitive to warnings, creating a need for a TOR system that balances urgency with credibility to prevent "alarm fatigue" [14] or over-reaction. This presents a novel perspective in TOR design, as prior research has not considered how alert satura-

tion due to cultural context influences driver perception of TORs. This will vary widely across the world and may affect adoption of automated vehicles in unforeseen ways. One solution is to design TORs that dynamically reduce non-critical notifications, like routine vehicle updates, to minimise alert fatigue and ensure important information is conveyed. Furthermore, coinciding with previous research, our thematic analysis indicates that drivers in more hazardous environments, such as Israel, develop a higher threshold for identifying hazards, often recognising them only when they are particularly severe [148]. In cultures accustomed to high-risk situations, a TOR may be ignored if it does not align with the driver's personal risk assessment. To address both the altered alert sensitivity and high hazard threshold, Israeli participants proposed a **Risk-Aware TOR** by including a RiskMeter (see Figure 5.4 and Table 5.1)—an algorithm designed to differentiate between varying levels of risk and tailor alerts accordingly. This approach ensures that TORs are credible and effective, matching drivers' varying risk perceptions and preventing both overreaction and desensitisation. Reich et al. [126] introduced a similar concept, the Dynamic Risk Assessment (DRA) method, which leverages tactical situational knowledge to mimic human-like risk reasoning. While a RiskMeter may lead to better compliance and stress management following a TOR, the ethical dilemma of prioritising hazards persists, often illustrated by the *Trolley problem* [39]. For instance, should the system classify the risk of hitting a human as more critical than hitting an animal? Such decisions are not only complex but also influenced by personal beliefs, adding another layer of complexity.

Drivers from cultures with low adherence to traffic laws and/or altered alert sensitivity may circumvent TORs, posing significant challenges for TOR design and enforcement. These findings open critical, novel research avenues for Level 3 car development, highlighting the need for culturally adaptive solutions that cater for drivers' driving culture and alert sensitivity.

5.5.1 Designing culturally- and age-sensitive TORs

Despite the cultural and age differences observed in our findings, designers can still create TORs that generalise across cultures and age groups by integrating core commonalities while offering customisation to accommodate differences. A key concept across all groups and all NDRTs was the **Context-Aware TOR** (see Table 5.1), which not only issues a takeover command but also provides situational details like hazard type and location, enhancing situational awareness [92]. Contrary to prior research suggesting TOR adaptation based on NDRT sensory demands [111], our study—the first to conduct designs around multiple NDRTs—revealed a preference for a uniform context-aware TOR across all tasks involving a device, despite varying perceived workloads as indicated by NASA-TLX scores (see Figure 5.2). This preference likely reduces real-time interpretation, streamlining design toward a consistent TOR system rather than tailoring alerts for each NDRT. Another significant finding was the preference for the **Dual-TOR** (see Table 5.1), where personal devices such as smartphones or tablets are repurposed to issue

TORs alongside the vehicle's in-car system. This finding is particularly novel when contrasted with previous literature [64, 128, 92], which did not utilise settings that enabled participants to incorporate their own devices into TOR design. Our study highlights the growing role of these devices in enhancing user engagement and responsiveness, offering a fresh perspective on TOR customisation. Across the study, all participants preferred multi-modal alerts, with older drivers opting for tri-modal and younger drivers favouring bi-modal, both consistently incorporating the visual modality. This preference can be accommodated through binary customisation, i.e., allowing drivers to turn on/off a modality. However, based on our findings, we recommend always including the visual modality and not allowing single-modality alerts as they proved less effective than multi-modal ones [166, 122]. Additionally, ordinal customisation is suggested, enabling drivers to fine-tune the intensity of auditory, and vibrotactile cues based on their sensory preferences. For example, older drivers might benefit from stronger vibrotactile feedback to offset sensory decline.

5.6 Limitations and future work

While this study provides valuable insights into the design of Take Over Requests for drivers across two age groups, two cultural contexts and four NDRTs, several limitations should be acknowledged. First, the design study was conducted in a stationary car, which may not fully reflect the complexities of real-world driving. The design concepts we propose need to be empirically tested in a more realistic driving scenario to see how they perform in varied driving conditions and across other NDRTs, which is our next step. Future work should also evaluate the cognitive demands of these designs and explore potential information saturation or cognitive cost. This would further refine the TOR designs, and enhance driver safety and efficiency in real-world applications. Second, some participants preferred reading from a physical book rather than a smartphone, suggesting that TORs may need to be projected onto the book or integrated with AR (Augmented Reality) technology for the Dual-TOR design concept to be possible. Third, hazard criticality in the context of Level 3 can range between unexpected road events (higher criticality) to pre-planned take overs (lower criticality). In this study, the design was for non-critical hazards. We recommend evaluating these designs across different hazard criticality levels and iteratively optimising them as necessary. Fourth, future research could expand on this study by exploring additional cultural contexts to further understand how different cultural factors influence TOR perception and effectiveness. Fifth, ethical and legislative aspects were not the focus of these design sessions. The development of TORs that adapt to driver age or cultural background introduces complex ethical questions, such as the risk of discrimination or bias in system design. Addressing these concerns will be crucial for ensuring fairness and inclusivity in future implementations. Sixth, regional legislation significantly influences TOR deployment, and further research is needed to explore how these systems can adapt to different legal standards while ensuring compliance and effectiveness across regions.

5.7 Conclusion

We presented the first participatory design study on Take-Over Request interfaces for Level 3 autonomous vehicles, examining the interplay of age, cultural context, and non-driving-related task type. We conducted a cross-cultural study across the UK, where there is high compliance with traffic laws, and Israel, with lower compliance, with younger and older drivers. The results revealed both distinct and similar preferences across groups: older drivers favoured tri-modal TORs for their redundancy and sensory reinforcement, particularly valuing the suspension of NDRT activities, while younger drivers preferred gradual bi-modal TORs, allowing them to maintain control over NDRT interruptions. This finding strongly suggests that existing behavioural trends will carry over into the realm of Level 3 autonomous cars. Younger drivers' inclination toward riskier behaviour and their expectation for systems that support multitasking remain consistent, even in the context of advanced vehicle technology. Israeli participants, influenced by regional conflict and low compliance with traffic laws, introduced the idea of a Risk-Aware TOR, which dynamically adjusts based on hazard criticality. This finding probes an unexplored area of TOR design in regions of instability, where alert sensitivity and risk perception may be altered. A universal preference across all groups was for unified TORs —conveying take-over command, hazard location, and type—across all NDRTs while repurposing NDRT-devices, such as phones, for delivering the take over command. This challenges prior suggestions to adapt TORs based on task demands, underscoring the importance of consistency in design. This study lays the foundation for understanding how cultural and age-related factors influence the intricacies of TOR design in Level 3 autonomous vehicles. Our findings should prove valuable for regulators and designers involved in shaping TOR features.

Chapter 6

Examining the Role of NDRT Suspension in Facilitating Safe Control Transitions in Level 3 Cars (conditionally accepted at AutomotiveUI '26)

6.1 Introduction

Autonomous vehicles are poised to revolutionize transportation by enhancing safety, reducing energy consumption, and increasing accessibility, while reshaping urban infrastructure and travel costs [12]. Among the SAE-defined levels of automation (from Level 0: no automation to Level 5: full automation) [1], Level 3 marks a pivotal advancement. These vehicles permit drivers to disengage from driving and engage in Non-Driving-Related Tasks (NDRTs), yet still require them to resume control upon a Take-Over Request (TOR) when the system encounters issues. This introduces a challenge: as drivers become less involved in the driving task, their reaction times increase and situational awareness declines—particularly when occupied with visually demanding NDRTs [43, 168, 142]. With Level 3 systems already deployed in countries like the US and Germany [53], optimizing TOR design has become a critical priority for ensuring safe and effective control transitions.

The ability to safely resume control following a TOR is closely linked to executive function—namely, the driver's ability to inhibit ongoing activity, shift attention, and update their mental model of the driving environment [117]. These cognitive functions are loaded during engagement with demanding NDRTs such as reading, video watching, or typing [116, 140]. This challenge is true for drivers of all ages but is amplified in older drivers due to age-related declines in attention and processing speed [117]. Research shows that they struggle more during takeovers than their younger counterparts, espe-

cially when engaged in demanding NDRTs [117, 30, 165]. Compared to younger drivers, older drivers exhibit fewer glances toward hazard-relevant areas and shorter mirror fixation durations when engaged in NDRTs [116]. They also exhibit longer reaction and takeover times, maintain shorter following distances [91], and show decreased performance when switching between NDRTs [158]. However, "older drivers" are often treated as a homogeneous group, masking potential differences in performance, needs, and risk profiles within this diverse population. Few studies have investigated if there are differences between drivers in the 60–69 age range and those over 70 years. The former often retain better functional capacity and self-regulatory strategies, whereas the latter are more likely to show pronounced cognitive and physical decline [141, 140, 99], suggesting that the different subgroups may require different support for TORs.

. Comparing these two subgroups allows researchers to identify whether TOR support strategies benefit older drivers uniformly or differently depending on their stage of ageing.

In 2013 in the UK, over 4 million drivers were aged 70+, with projections showing over 25% of EU drivers to be 70+ by 2050 [131]. In the US, those over 65 will account for a quarter of driver fatalities by 2030 [117]. The rising numbers of older drivers and the control transition challenge introduced by Level 3 cars make tailoring TOR design for older drivers imperative to ensure smooth and safe deployment.

Younger drivers, meanwhile, frequently use smartphones while driving despite legal restrictions, engaging in visually and cognitively demanding tasks such as texting and browsing [79]. These behaviours impair situational awareness and reaction time, increasing crash risk even among younger cohorts [108]. While Level 3 automation provides a safer context by legally allowing such activities during automated driving, younger drivers—like their older counterparts—can still experience delayed responses and degraded situational awareness during control transitions, particularly when re-engaging from demanding NDRTs. Despite these limitations, both older and younger drivers are keen to engage in a variety of NDRTs while using Level 3 automated vehicles [92, 141], underscoring the need for tailored design interventions.

To mitigate the control transition challenge in Level 3, prior research has explored various TORs—such as modality-based cues [122, 66, 67], and multi-sensory feedback from the in-vehicle information system (IVIS) [67, 140]. Most research on TORs in Level 3 vehicles focuses on cue timing and modality, largely overlooking whether, during a take-over, the NDRT should be suspended (i.e., the complete cessation of the NDRT by the system during the take-over time window, preventing the driver from continuing the NDRT until the takeover manoeuvrer is over) —an important but underexplored question. **Testing task suspension is important because it offers a system-driven intervention that may alleviate the cognitive burden of self-regulated disengagement, especially in high-demand or time-critical situations.** Some studies suggest its potential: older drivers have proposed NDRT suspension to improve safety [142], and prior

work showed that issuing the TOR via the NDRT device enhanced takeover performance, especially for drivers aged 70+ [140]. However, neither approach tested task cessation directly.

The study presented in this paper is the first to investigate the effects of *complete NDRT suspension*—where the secondary task is forcibly blocked—at the moment a TOR is issued in a Level 3 automated driving context. Using a driving simulator, participants engaged in two types of NDRTs on a smartphone—one with higher and one with lower cognitive demand—while responding to TORs. In this study we: **(1)** Evaluated whether suspending the NDRT during a TOR improves takeover reaction time, hazard perception and user acceptance across older and younger drivers, compared to a condition where the NDRT remains active. **(2)** Examined age-related differences in TOR performance and acceptance across three driver cohorts—young (22–32), younger-older (60–69), and older-old (70+)—in both conditions: when the NDRT is suspended and when it remains active. **(3)** Assessed whether the impact of NDRT suspension on performance and acceptance generalizes across both low- and high cognitive-demand NDRTs.

Contribution This paper demonstrates that *complete NDRT suspension at TOR onset* reduces takeover reaction time across both low- and high-demand tasks and across three age groups (22–32, 60–69, 70+). Crucially, it shows that NDRT suspension eliminates age-related disparities observed in the non-suspended condition, with the largest gains in drivers aged 70+. By jointly evaluating reaction time, hazard perception, workload, and user acceptance, this work establishes both the benefits and limitations of task suspension as a practical design strategy for Level 3 vehicles.

Novelty This work is the first to introduce *task suspension as a system-level manipulation* and demonstrates that its effectiveness depends on task demand and age subgroup. It identifies a non-linear age effect, where drivers aged 70+ benefit most, the 60–69 group shows limited gains, and younger drivers benefit moderately, challenging the treatment of older drivers as a homogeneous group.

Significance The findings show that while suspension improves reaction time across conditions, it enhances hazard perception only for low-demand tasks and increases workload under high-demand tasks. This highlights that *task management, not just TOR design, is critical*, and supports the development of adaptive, context-aware TOR systems that account for cognitive demand and driver heterogeneity.

6.2 Related Work

6.2.1 Non-Driving Related Task Interruption in Level 3 Cars

Engaging in NDRTs during Level 3 driving impairs drivers' ability to respond to Takeover Requests (TORs) [154, 117, 150]. A recent meta-analysis of 48 studies confirms that en-

gaging in non-driving related tasks results in degraded takeover performance, particularly if it has overlapping resource demands with the driving task [154]. Task characteristics play a critical role in shaping takeover performance [168]. Muller et al. [108] compared NDRTs for reading, listening, and watching movies and found that different NDRTs had significant differences in psychological workload, and NDRTS associated with high workload (such as reading) would lead to longer reaction time. To mitigate the effect of NDRT engagement on takeover performance, researchers have proposed various NDRT interruption strategies. For example, Janssen et al. [70] outlined multiple stages of disengagement, including interleaving, forced interruption, and task suspension. Janssen et al. [70] and Glenn et al. [161] emphasize that even non-demanding tasks can leave behind cognitive residue, undermining driver ability to re-engage promptly with the driving task [70]. These studies [70, 161], however, did not evaluate task suspension vs. self-initiated disengagement and did not compare its impact across varying NDRT cognitive demands or age groups.

In self-initiated disengagement, drivers must recognise the TOR, decide to stop the ongoing NDRT, and redirect their attention to the driving environment while the secondary task remains available, requiring them to manage the transition between two competing tasks [70, 129]. By contrast, task suspension involves the system actively pausing the NDRT at TOR onset, removing the need for self-regulated disengagement and potentially facilitating a faster reallocation of attention to the road. Such support may be particularly beneficial in time-critical situations and for populations with reduced cognitive flexibility, such as older adults [7, 116, 77, 142]. This rationale is further supported by findings from a participatory design study, in which older drivers themselves proposed suspension across four NDRT types as a means of improving takeover readiness [142]. However, suspension may still incur task-switching costs or generate frustration [132, 161]. Therefore, it is essential to evaluate whether its benefits outweigh these potential drawbacks across different age groups and task demands. Our work expands on previous literature by systematically evaluating the effects of suspending ongoing NDRTs at the moment of a TOR, using both high- and low-demand tasks, and comparing responses across young (22-32), younger-old (60-69) and older-old (70+) drivers. This provides the first empirical evidence for the feasibility and limitations of NDRT suspension as an age- and task-sensitive intervention in Level 3 vehicles.

6.2.2 Age-Related Differences in Takeover Performance

Compared to younger drivers, older drivers demonstrate delayed and less effective responses to take-over requests, particularly when engaged in non-driving related tasks [44, 117, 30]. These performance differences are attributed to age-related declines in attention shifting, processing speed, and working memory [7, 91]. While older drivers' performance may be hindered by these age-related declines, younger drivers—despite intact cognitive and physical capabilities—often exhibit riskier driving behaviours [147] and a higher propensity to engage in NDRTs [73], which can also interfere with takeover per-

formance. Older drivers show slower reactions, reduced hazard awareness, and greater sensitivity to cognitively demanding NDRTs [44, 117, 30]. For instance, Yoon et al. [165] observed that TOR reaction times were significantly longer when participants engaged in smartphone-based NDRTs, and older adults were more affected than younger ones. Eye-tracking data from Peng et al. [116] revealed that older drivers engaged in fewer gaze transitions between mirrors and hazard-relevant areas, indicating lower situational awareness during automation. Age also interacts with NDRT type and complexity. Wu et al. [158] found that older adults (mean age 64.3) exhibited greater performance degradation when switching between multiple NDRTs, compared to younger adults. Similarly, Ebnali et al. [30] showed that cognitively demanding NDRTs disproportionately impaired takeover quality in drivers aged 65 and older. These findings indicate that both the nature of the NDRT and the cognitive demands of control resumption pose greater risks for older users. Despite these impairments, older drivers remain willing to engage in various NDRTs while the vehicle is automated [92, 141]. At the same time, they report higher awareness of their limitations and support interface designs that suspend ongoing tasks upon TOR issuance [92, 141, 142]. Importantly, the older adult group is not homogeneous. While age 60 is widely used as a lower threshold for "older driver" classification [93, 7, 87], performance variations increase markedly with age [141, 140]. In a comparative study, Li et al. [93] demonstrated that drivers aged 70 and above performed worse in takeover situations than those aged 60–69, with significantly slower reaction times and poorer completion rates. Road safety data corroborates this divide: individuals aged 70+ have higher rates of traffic-related fatalities and severe injuries per mile driven compared to younger cohorts [25].

In summary, age-related differences in takeover performance are well established, with drivers aged 70+ showing greater impairments, particularly under high-demand NDRTs. These findings highlight the need for TOR designs that account for variability across age groups.

In this paper, we address the following research questions:

•**RQ1** How does complete suspension of the NDRT during a TOR affect the takeover reaction time across two older driver and one younger driver cohorts, considering two NDRTs with varying cognitive demand?

• **RQ2** How does complete suspension of the NDRT during a TOR affect hazard perception across two older driver and one younger driver cohorts, considering two NDRTs with varying cognitive demand?

•**RQ3** Does complete suspension of the NDRT enhance TOR acceptance and detection across two older driver and one younger driver cohorts, considering two NDRTs with varying levels of cognitive demand?

6.3 Methodology

6.3.1 Participants

We recruited 40 licensed drivers divided into three age groups: 22–32 (M = 25.1, SD = 3.2; 6 females, 10 males), 60–69 (M = 64.5, SD = 3.5; 5 females, 7 males), and 70+ (M = 75.5, SD = 3.7; 6 females, 6 males). The subdivision of the older cohort followed established classifications in the automated driving literature, where age 60 is widely adopted as the lower threshold for "older driver" classification [7, 87, 93]. The subdivision at 70 is justified by evidence that drivers aged 70+ perform significantly worse in takeover situations than those aged 60–69 [93]. Recruitment was done using university mailing lists and local community outreach. All participants had been active drivers for at least three years and reported normal or corrected-to-normal vision and hearing. None had prior experience with automated vehicles.

Participants completed two sessions lasting approximately one hour each and received £10 compensation. Cognitive screening was conducted using the Montreal Cognitive Assessment (MoCA) [109], with all participants scoring above the threshold for intact cognition (M = 28.1, SD = 1.5). Two participants completed only one session, resulting in partial data. Ethical approval was obtained from the institutional ethics committee.

6.3.2 Study Design

We employed a 2 (NDRT cognitive demand: higher vs. lower) × 2 (NDRT status: active vs. suspended) × 3 (age group: 22–32, 60–69, 70+) mixed factorial design. Cognitive demand and NDRT status were within-subject factors, while age group was a between-subject factor.

The study builds on prior experimental paradigm by Srour-Zreik et al. [140], adapting the same task structure and takeover procedure. The primary extension is the manipulation of **NDRT status at TOR onset**, where the NDRT was either suspended or remained active during the takeover window.

Higher-demand NDRTs consisted of reading comprehension tasks with multiple-choice questions, while lower-demand NDRTs involved passive video viewing. The combination of factors resulted in four experimental conditions: **(1)** higher-demand NDRT with task suspension (reading comprehension, Set 1), **(2)** lower-demand NDRT with task suspension (video, Clip 1), **(3)** higher-demand NDRT with task active (reading comprehension, Set 2), and **(4)** lower-demand NDRT with task active (video, Clip 2). Participants completed all conditions in a counterbalanced order.

The following dependent variables were collected: (1) **Hazard perception:** (a) time (in seconds) from TOR onset to the first fixation (>100 ms) on the AOI, and (b) total number of fixations (>100 ms) on the AOI, both measured and collected using an eye tracker. (2) **Takeover performance:** reaction time (in seconds) from TOR onset to manual takeover (button press on the steering wheel, collected through Psychopy software). (3) **Mental workload:** assessed using the NASA-TLX (6 items, 0–10 scale). (4) **User acceptance of**

TOR: measured using the UAATT questionnaire (9 items, -2 to 2 scale). (5) **Ease of TOR detection:** measured using the Single Ease Questionnaire (SEQ; single item, 1–7 scale).

6.3.3 Apparatus

The experimental setup builds on the simulator-based setup by Srour-Zreik et al. [140].

Driving Environment

Participants completed the experiment in a fixed-base driving simulator displaying pre-recorded urban driving scenarios on a 27-inch monitor (Asus VX279) positioned approximately 1 metre from the participant. The experiment was implemented in PsychoPy v2021.2.3 [115]. Driving footage was recorded from a driver's perspective using a GoPro Max camera mounted on a Citroën C3 [140]. The recordings were edited into four scenarios (11–13 minutes each), incorporating naturally occurring hazards such as pedestrians crossing, cyclists, vehicles entering the lane, and roadworks. Six hazards per scenario were selected for TOR issuance based on established hazard perception criteria [28, 168, 149]. Each hazard was associated with a predefined Area of Interest (AOI) used for gaze-based analysis.

Non-Driving Related Tasks (NDRTs)

Participants engaged in visual-cognitive NDRTs presented on a Google Pixel 7 smartphone via a custom-built Flutter application. Two levels of cognitive demand were implemented: **(1)** Lower-demand:passive video viewing (e.g., *Tom and Jerry*, *Mr. Bean*) and **(2)** Higher-demand: reading comprehension tasks with multiple-choice questions. Two equivalent versions of each task type (two videos, two reading sets) were used and distributed across conditions. Task selection was based on prior validation showing significantly higher perceived workload for reading compared to video viewing [108, 140].

Take-Over Request (TOR) Design

TORs were tri-modal (visual, auditory, vibrotactile), consistent with prior recommendations [142, 140]. TORs were presented simultaneously via the in-vehicle information system (IVIS) and the smartphone (dual TOR). The IVIS presented a tri-modal TOR: visual heads-up display (HUD) message indicating the hazard type and location, accompanied by a spoken auditory alert delivered via external speakers and vibrotactile feedback through seven actuators embedded in the simulator seat. In parallel, the smartphone presented a visual takeover message with a dynamic countdown (see Figure 6.1), alongside vibrotactile feedback via the device's internal actuator. All TORs were issued within a fixed 10-second response window, consistent with prior literature [103, 140].

Eye movements were recorded using Tobii Pro Glasses 2, and fixations exceeding 100 ms were extracted using Tobii Pro Lab software and mapped onto predefined Areas of Interest (AOIs) to assess hazard perception.

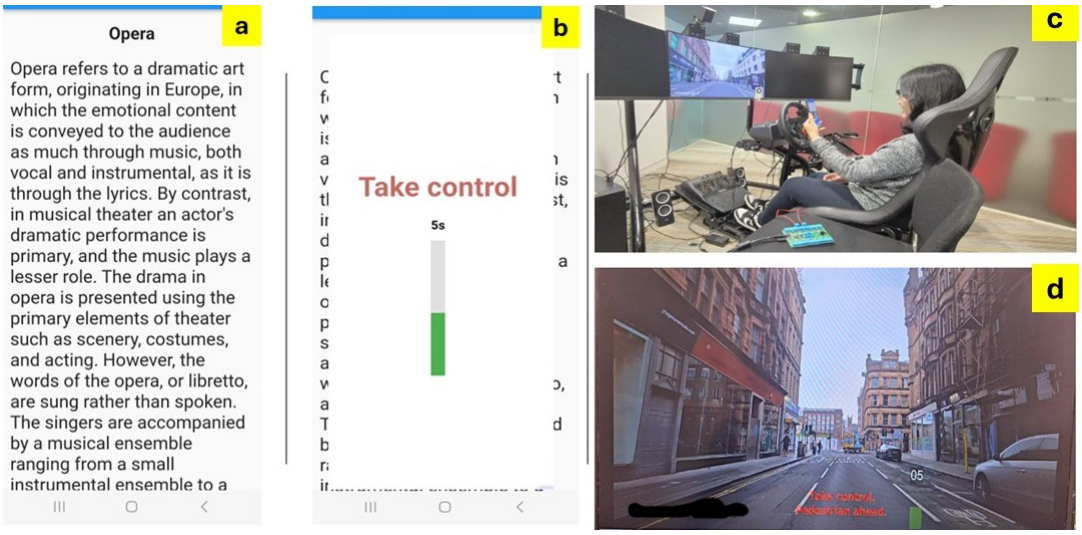


Figure 6.1: NDRTs presented on the smartphone: (a) reading comprehension NDRT, (b) reading comprehension NDRT suspended by a TOR, (c) a participant from age group 22-32 wearing the eye-tracker and engaging with an NDRT on the smartphone while Level 3 automation is on, (d) the visual component of the TOR issued on the head-up display (the TOR message says "Take control. Pedestrian ahead")

6.4 Procedure

After providing consent, participants completed a demographic questionnaire and a cognitive screening (MoCA) administered by a certified experimenter. They were then introduced to the driving simulator and completed a brief training session to familiarise themselves with the environment, NDRTs, and Take-Over Requests (TORs).

Adapting the protocol by Srour-Zreik et. al [140], the experiment was conducted across two sessions (approximately one hour each). In each condition, participants engaged in an NDRT on a smartphone while viewing a pre-recorded Level 3 driving scenario. They were instructed that active monitoring was not required unless a TOR was issued and were encouraged to focus on the NDRT.

During each scenario, six TORs were triggered on both the in-vehicle display and the smartphone. Depending on condition, the NDRT was either suspended or remained active. Upon receiving a TOR, participants were instructed to press a steering wheel button as quickly as possible to simulate takeover. There was no other input from the steering wheel or pedals since this was a pre-recorded drive. After each event, automation resumed and participants returned to the NDRT. Conditions were counterbalanced across participants.

Following each condition, participants completed NASA-TLX, UAATT, and SEQ questionnaires via Qualtrics to assess workload, user acceptance, and ease of TOR detection. Participants received £10 upon completing both sessions.

6.5 Results

The results were analysed using R Studio 2022.07.2 Build 576 using the lme4 [10], lmerTest [83] and report [97] Rpackages.

6.5.1 Take-Over Performance

A Linear Mixed Model examined effects of Condition and Age Group on Reaction Time, with random intercepts for participants and TOR order. Despite non-normal residuals, model assumptions were deemed acceptable given the robustness of mixed-effects models [130].

Watching a Video NDRT. Six data points (i.e., six TOR incidents) were omitted from the analysis due to instances where participants either dropped the smartphone or failed to adhere to study instructions. We found a significant main effect of NDRT status with a large effect size *($F(1,10.77)=92.76$, $p<0.0001$; $\eta^2=0.86$)* and a significant interaction between NDRT status and Age Group with a medium effect size *($F(2,334.84)=12.01$, $p<0.0001$; $\eta^2=0.07$)*. Overall, the NDRT-suspended condition led to a significantly shorter reaction time *(Estimated Mean (EM)=1.84 sec, SE=0.13 sec)* than the NDRT-active condition *(EM=2.82 sec, SE=0.13 sec)*. *Post hoc* analyses with a Bonferroni adjustment revealed that reaction times of all groups were shorter when the NDRT was suspended compared to when NDRT remained active, see Figure 6.2. Additionally, in the NDRT-active condition, the 70+ age group exhibited significantly longer reaction times (EM=3.27 sec, SE = 0.21 sec; p=0.015) than the 60–69 group *(EM=2.47 sec, SE = 0.2 sec)*. This difference was not observed when the NDRT was suspended *(p>0.05)*, see Figure 6.2 (red annotation) and Table 6.1.

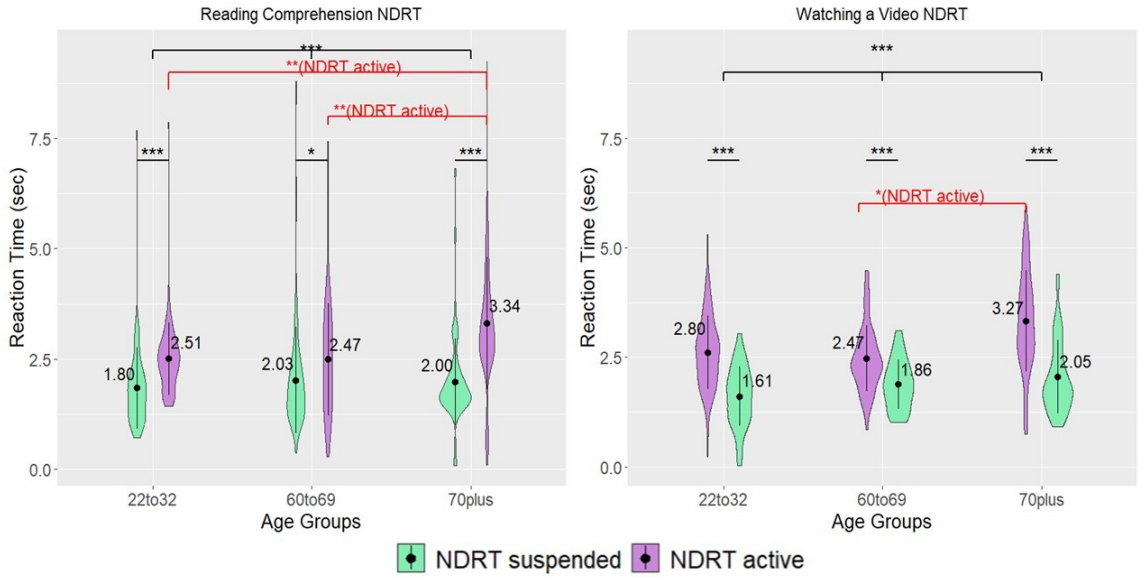


Figure 6.2: The interaction between Age Group and Condition for each NDRT for the dependent variable Reaction Time. The dots represent estimated mean values. *** - P<0.001, ** - P<0.01, *- P<0.05.

Reading Comprehension NDRT. We found a significant main effect of Condition with a large effect size *(F(1,11.7)=58.78, p<0.0001; η^2=0.83)* and a significant interaction between Condition and Age Group with a medium effect size *(F(2,312.95)=12.34, p<0.0001; η^2=0.07*). Overall, the NDRT-suspended condition led to a significantly shorter reaction time *(Estimated Mean (EM)=1.8 sec, SE=0.15 sec)* than the NDRT-active condition *(EM=2.74 sec, SE=0.14 sec)*. *Post hoc* analyses with a Bonferroni adjustment revealed that reaction times of all groups were shorter when the NDRT was suspended compared to when it remained active, see Figure 6.2. Additionally, in the NDRT-active condition, the 70+ age group exhibited significantly longer reaction times (EM=3.34 sec, SE = 0.24 sec; p<0.01) than the 60–69 group *(EM=2.47 sec, SE = 0.24 sec)* and than the 22-32 group *(EM=2.51 sec, SE = 0.22 sec; p=0.03)*. This difference was not observed when the NDRT was suspended *(p>0.05)*, see Figure 6.2 (red annotation) and Table 6.1.

6.5.2 NASA-TLX Overall Workload

We conducted an ANOVA on a Linear Mixed Model to predict the main effects of Condition and Age Group on overall NASA-TLX scores with a random intercept for each participant.

Watching a Video NDRT. The model did not yield any significant main effects and no significant interaction between Condition and Age group was found *(F(2,24)=0.62, p=0.54)*.

Reading Comprehension NDRT. We found a significant main effect of Condition *(F(1,29.5) = 4.47, p = 0.04; η^2 =0.13)*. The NDRT-suspended condition yielded significantly higher overall NASA-TLX scores *(M=25, SD=8.86)* than NDRT-active *(M=21.9, SD=10.6)*. The interaction between Age group and Condition was not significant *(F(2,30.72) = 0.85, p =0.43)*.

6.5.3 Hazard Perception

Each analysis began with a model that included Age Group and Condition as fixed effects, and participant and TOR order as random effects. A backward elimination approach was used to refine the model, removing all non-significant interaction terms. For any significant fixed effects with more than two levels, Bonferroni-adjusted *post hoc* pairwise comparisons were conducted.

Watching a video NDRT

Number of Fixations (>100 ms) on AOI The Number of Fixations on the AOI had a Poisson distribution. We conducted an ANOVA on a Poisson regression model with a log link function within the Generalised Linear Mixed Models (GLMM) framework. The model's total explanatory power was moderate (conditional R2 = 0.3) [1]. Within this

[1]Whilst it is difficult to calculate the actual R2 value for generalised linear models, the theoretical value was calculated using the MuMIn package [9]. R2 values above 0.2 are generally

Table 6.1: Effect of NDRT Suspension on Reaction Time by NDRT Type.

NDRT Type	Main Effect	Statistic and Value	Post Hoc	Takeaway
Video	Condition	$F(1, 10.77) = 92.76$ $p < .0001$, $\eta^2 = .86$	Suspended (EM = 1.84, SE = 0.13) $<$ Active (EM = 2.82, SE = 0.13) for all age groups.	Suspension reduces RT across all age groups
	Cond. × Age	$F(2, 334.84) = 12.01$ $p < .0001$, $\eta^2 = .07$	70plus (EM=3.27, SE = 0.21) $>$ 60–69 (EM=2.47, SE = 0.20) in Active only ($p = 0.015$); no difference in Suspended	Suspension eliminates age-group disparity
Reading	Condition	$F(1, 11.7) = 58.78$ $p < .0001$, $\eta^2 = .83$	Suspended (EM = 1.80, SE = 0.15) $<$ Active (EM=2.74, SE = 0.14) for all age groups.	Clear benefit of suspension across all groups
	Cond. × Age	$F(2, 312.95) = 12.34$ $p < .0001$, $\eta^2 = .07$	70plus (EM=3.34, SE=0.24) $>$ 60–69 (EM=2.47, SE=0.24; $p < .01$) and $>$ 22–32 (EM=2.51, SE=0.22; $p = 0.03$) in Active only, No difference in Suspended	Suspension eliminates age-group disparity

model a significant main effect of Condition ($\chi^2(1) = 6.96$, $p<0.01$) was found, NDRT-suspension yielded a higher number of fixations on the AOI (*Estimated Mean (EM)=3.84, SE=0.35*) than NDRT-active condition (*EM=2.66, SE=0.27*). Age Group did not show a significant effect on the number of fixations ($\chi^2(2) = 1.2$, $p=0.53$) and no interaction between Age Group and Condition was found either ($\chi^2(2) = 0.42$, $p=0.8$).

Time to First Fixation on AOI (TTFF) A Linear Mixed Model ANOVA was conducted; normality was confirmed via Shapiro–Wilk. The interaction between Age Group and Condition was significant ($F(2,305.2)= 7.46$, $p<0.001$, $\eta^2 =0.05$). *Post hoc* analyses with a Bonferroni adjustment revealed that for participants in the age group 60-69, no significant difference was observed between conditions ($p=0.6$). However, in the age groups 70plus and 22-32, a significant difference emerged between conditions ($p<0.05$). Both groups had significantly lower TTFF when the NDRT was suspended. See Figure 6.3. Additionally, in the NDRT-active condition, both the 70plus and 22-32 age groups exhibited significantly longer TTFF than the 60–69 group. This difference was not observed when the NDRT was suspended *($p>0.05$)*. See Figure 6.3.

considered good [101].

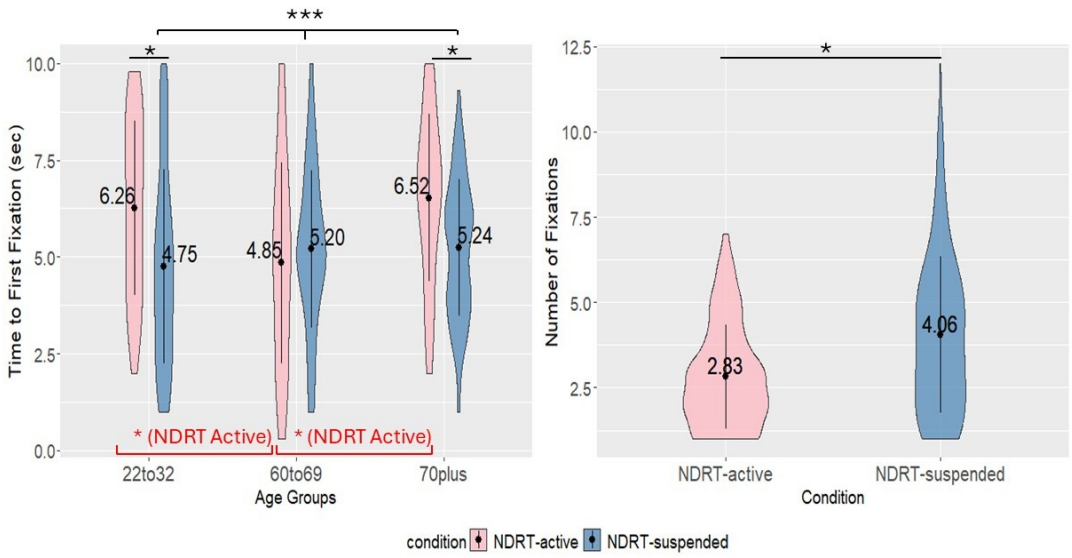


Figure 6.3: Time to First Fixation and Number of Fixations on AOI for Watching a Video NDRT.

Reading Comprehension NDRT

Number of Fixations (>100 ms) on AOI The dependent variable Number of Fixations on the AOI had a Poisson distribution. We conducted an ANOVA on a Poisson regression model with a log link function within the Generalised Linear Mixed Models (GLMM) framework. Results revealed a significant main effect of age group, ($\chi^2(2) = 6.49$, $p = .039$). *Post hoc* comparisons showed that participants aged 60–69 had significantly higher fixation counts than those aged 70+ (ΔEM=0.51, $p = .001$). No significant main effect of condition, ($\chi^2(1) = 0.79$, $p = .374$). The interaction between age group and condition was also not significant, ($\chi^2(2) = 0.30$, $p = .860$).

Time to First Fixation on AOI (TTFF) We conducted an ANOVA on a linear regression model within the Linear Mixed Models (LMM) framework with a normal distribution for the dependent variable. Normality of the data was confirmed via Shapiro-Wilk test. The model did not yield any significant results.

6.5.4 User Acceptance of Advanced Transport Telematics Questionnaire (UAATT)

The Usefulness and Satisfaction scores were derived as instructed in [146]. A reliability analysis indicated satisfactory internal consistency, with a Cronbach's alpha of 0.81 for Satisfaction items and 0.89 for Usefulness items [146]. Due to non-normality (Shapiro–Wilk), an ART ANOVA with participant as a random intercept was conducted to examine effects of Condition and Age Group on Usefulness and Satisfaction.

Watching a Video NDRT. For the dependent variable Usefulness, the main effect of Condition was statistically significant *($F(1,34.94) = 22.61$, $p < .001$; $\eta^2 = 0.39$)*. The NDRT-

suspended condition yielded significantly higher Usefulness scores *(M = 1.7, SD = 0.41)* than the NDRT-active condition *(M = 1.25, SD = 0.77)*. A significant main effect of Age Group was also observed *(F(2,33.61) = 3.42, p = .045; η^2 = 0.17)*, while the interaction between Age Group and Condition was not significant *(F(2,35.04) = 1.74, p = .191)*.

Reading Comprehension NDRT. For Satisfaction and Usefulness both models did not yield any significant main effects.

6.5.5 Single Ease Questionnaire

Due to non-normality (Shapiro–Wilk), an ART ANOVA with participant as a random intercept was conducted to examine effects of Condition and Age Group on Ease of Detection.

Watching a Video NDRT. We found a significant main effect of Condition *(F(1,38.83) = 5.27, p = 0.027; η^2 =0.2)*. The NDRT-suspended condition yielded significantly higher Ease of Detection scores *(M=6.25, SD=1.25)* than NDRT-active *(M=5.59, SD=1.58)*. The interaction between Age group and Condition was not significant *(F(2,37) = 0.28, p =0.7)*.

Reading Comprehension NDRT. We found a significant main effect of Condition *(F(1,33.88) = 8.66, p < 0.01; η^2 =0.12)*. The NDRT-suspended condition yielded significantly higher Ease of Detection scores *(M=6.48, SD=0.9)* than NDRT-active *(M=5.76, SD=1.38)*. The interaction between Age group and Condition was not significant *(F(2,31.97) = 2.21, p =0.12)*.

6.6 Discussion

The discussion is organised around the three research questions:

- **RQ1 How does complete suspension of the NDRT during a TOR affect the takeover reaction time across two older driver cohorts and one younger driver cohort, considering two NDRTs with varying cognitive demand?**

Watching a Video and Reading Comprehension NDRTs. Overall, results demonstrated a clear benefit of suspending the NDRT at the moment of a TOR.

For both the lower cognitive demand task (watching a video) and the higher cognitive demand task (reading comprehension), participants exhibited significantly faster reaction times when the NDRT was suspended compared to when it remained active.

Suspending the NDRT significantly reduced reaction times across both low- (video) and high-demand (reading) tasks, consistent across age groups. This extends prior work [140], showing that suspension improves performance even under high cognitive demand, effectively removing the moderating effect of task type.

Our results here demonstrate that suspending the NDRT significantly improves reaction times across both lower- and higher cognitive demand tasks, eliminating the moderating effect of task demand. While prior research has linked the cognitive load of an

NDRT to takeover performance [108], our findings suggest that task suspension can neutralize this impact. Notably, the most substantial improvements were observed in drivers aged 70+, indicating that suspension may effectively reduce age-related delays in taking control of a vehicle. Overall, task suspension emerged as a strategy for lowering takeover reaction times in Level 3 automated vehicles.

While the reduction in reaction time was significant across all age groups, drivers aged 70+ showed the greatest benefit. In the higher cognitive demand NDRT, their reaction time decreased by 1.34 seconds, compared to 0.44 seconds for the 60–69 group and 0.71 seconds for the 22–32 group. In the lower cognitive demand NDRT, the reduction was 1.22 seconds for the 70+ group, compared to 0.61 seconds for the 60–69 group and 1.19 seconds for the 22–32 group. According to Li et al. [93], drivers aged 70 and older were slower to respond to and complete takeover requests than those aged 60-69. The results from this study suggest that these reaction times can be improved by suspending the NDRT during the TOR.

The interaction between age group and condition revealed that NDRT suspension moderated the impact of age. For the lower cognitive demand NDRT (video watching), the 70+ group was significantly slower than the 60–69 group but not the younger group in the NDRT-active condition (see Figure 6.2 and Table 6.1). This may reflect distinct challenges: while 70+ drivers show age-related slowing, younger drivers may have been more immersed or complacent [94], delaying their responses. In contrast, the 60–69 group may still retain better attentional flexibility, enabling quicker disengagement. Task suspension eliminated these age-group disparities. For the higher cognitive demand NDRT (reading), the 70+ group was significantly slower than both the 60–69 and younger groups in the NDRT-active condition (see Figure 6.2 and Table 6.1), likely due to compounded effects of cognitive ageing. Reading requires sustained attention and working memory—capacities that decline more sharply in the 70+ group—making it harder for them to disengage. In contrast, younger drivers and the 60–69 group likely retained more cognitive flexibility and inhibitory control, enabling quicker transitions back to the driving task. Here too, task suspension eliminated age-group disparity.

- **RQ2 How does complete suspension of the NDRT during a TOR affect hazard perception across two older driver cohorts and one younger driver cohort, considering two NDRTs with varying cognitive demand?**

Reading Comprehension (higher cognitive demand NDRT). Suspending the higher cognitive demand reading task during a TOR did not improve hazard perception in any age group. This lack of benefit can be attributed to the complex cognitive processes involved in reading comprehension, such as sustained attention and working memory [108]. Abruptly interrupting such tasks may cause attentional inertia and cognitive residue, which hinder effective re-engagement with the driving environment [129, 70]. Supporting this interpretation, the NASA-TLX results showed that suspending the reading task significantly increased subjective workload, with higher overall scores in the suspended condition compared to when the task remained active ($M = 25$ *vs.* $M = 21.9$, $p =$

.04). Notably, effort and mental demand were the most affected subscales, indicating that the interruption itself introduced cognitive strain rather than relieving it. The absence of a significant interaction with age suggests that these effects were consistent across age groups. These findings reinforce the notion that deep cognitive engagement imposes universal switching costs, as evidenced by prior work [161]. Notably, no improvements were observed in key hazard perception metrics—namely, the number of hazard fixations and time-to-first fixation—underscoring the limitations of interface-based interventions under high cognitive load. Even when TORs are carefully designed, they may be insufficient if drivers remain mentally anchored in demanding NDRTs. This aligns with findings by Srour-Zreik et al. [140], who reported that task interruption, rather than suspension, similarly failed to improve takeover performance for older drivers engaged in a higher cognitive demand Reading Comprehension NDRT. Together, these results suggest that task suspension alone may not be a sufficient strategy to enhance hazard perception in cognitively intensive scenarios.

Watching a Video (lower cognitive demand NDRT). In contrast, suspending the lower-cognitive demand video task yielded improvements in hazard perception across all age groups. The passive nature of the task [169] likely made it easier to disengage from, reducing dual-task interference and facilitating quicker attentional shifts to the driving environment. Notably, the NASA-TLX model for the video condition revealed no significant effects making it a low-cost task to interrupt.

The improvement in hazard perception can also be interpreted through the lens of arousal regulation. Watching a video during automated driving may induce a state of under-load [169]. According to the Yerkes-Dodson law, moderate increases in arousal can improve performance in tasks requiring attention [35]. We propose that suspending the video acted as an alerting cue, re-engaging participants cognitively and elevating them to a more optimal arousal state, thus enhancing their hazard perception. This aligns with prior work showing that small increases in workload can be beneficial when baseline engagement is low [153].

Notably however, in terms of time to first fixation (TTFF), in contrast to both the younger (22–32) and older (70+) groups who showed marked improvements, drivers aged 60–69 did not significantly benefit from the NDRT suspension. This group may have greater cognitive resilience than the 70+ group, while at the same time showing more cautious behaviour than the youngest group. As a result, they may already perform reasonably well without the NDRT suspension, thus leaving less room for improvement. These results highlight the need for age-sensitive adaptations in vehicle interface design, as task suspension may not offer uniform benefits across older driver subgroups.

- **RQ3 Does complete suspension of the NDRT enhance TOR acceptance and detection across the cohorts, considering two NDRTs with varying levels of cognitive demand?**

Watching a Video (Lower cognitive demand NDRT). Participants rated the suspen-

sion of the NDRT as more useful than maintaining task activity during the TOR. This aligns with the design recommendations by older drivers in [142]. In contrast, younger participants did not advocate for task suspension, preferring instead a gradual TOR [142]; however, when tested, they still found this design useful for lower cognitive-demand tasks. Satisfaction scores did not differ between NDRT statuses and were primarily associated with the likeability of the TOR itself, suggesting that further refinement of the suspension design could enhance user satisfaction.

Reading Comprehension (Higher cognitive demand NDRT). For the higher cognitive demand reading task, NDRT suspension did not significantly affect Satisfaction or Usefulness, likely because the abrupt interruption of a cognitively demanding task introduced mental strain without offering a clear benefit, limiting users' perceived value of the intervention.

6.7 Limitations and Future Work

This study offers novel evidence on the impact of NDRT suspension during Level 3 takeovers, but several limitations should be noted. First, only two visual–cognitive NDRTs (reading and video watching) were tested, limiting generalisability to other modalities. Suspension effects may differ for tasks with varying sensory or motor overlap with driving, such as auditory or physical activities. Second, in the high-demand reading task, abrupt suspension increased workload without improving hazard perception, indicating that more adaptive disengagement strategies—such as staggered cues or reorientation aids—should be explored. Third, the participant sample skewed toward highly educated individuals, which may limit applicability to broader populations with different cognitive profiles and technology familiarity. Fourth, the fixed 10-second TOR window and relatively frequent takeover events may not reflect naturalistic driving, where TORs occur less often and after longer disengagement periods. Future studies should vary timing and event frequency. Fifth, suspension was only implemented for NDRTs delivered via a connected smartphone. This limits applicability to non-system-controlled tasks (e.g., reading a physical book, talking to a passenger). Future work should explore interruption strategies for such activities. Finally, our findings support the need for context-aware systems that classify NDRTs by demand, modality, and permissibility, enabling personalised TOR responses tailored to driver characteristics and task context.

6.8 Conclusion and design implications

A key contribution of this work is the empirical validation of NDRT suspension, a participatory design concept proposed by older drivers [142]. While suspension consistently improved reaction times across all age groups and task types, its benefits did not extend to hazard perception under higher cognitive demand (reading comprehension), and it

increased subjective workload. This suggests that abrupt interruption of cognitively demanding tasks introduces switching costs that limit effective re-engagement, highlighting the importance of evaluating user-driven design concepts under realistic task conditions.

A non-linear age-related pattern emerged. Drivers aged 70+ showed the greatest improvements in reaction time, particularly under higher task demand, whereas the 60–69 group exhibited smaller and less consistent benefits, especially in hazard perception. Younger drivers also benefited, though to a lesser extent. These findings suggest that the 60–69 group may retain sufficient attentional flexibility to recover without intervention, while the 70+ group relies more heavily on system support. This challenges the treatment of older drivers as a homogeneous group and underscores the need for more fine-grained, behaviourally informed HMI design.

From a design perspective, task suspension is an effective and lightweight strategy for improving takeover speed, particularly for older drivers and low-demand tasks. However, it is insufficient for cognitively demanding tasks, where additional support—such as anticipatory cues or staged disengagement—may be required.

These findings also raise regulatory considerations. Although suspension improves reaction time, it does not mitigate the negative effects of high-demand NDRTs on hazard perception or workload. This suggests that certain NDRTs may require restriction or adaptive management in Level 3 systems until more effective interface solutions are developed.

Overall, this study advances understanding of both the potential and limitations of NDRT suspension, demonstrating that its effectiveness is shaped by task characteristics and driver heterogeneity. The results highlight a non-linear effect of age and call for adaptive, context-aware TOR designs that move beyond one-size-fits-all approaches to support safe and inclusive control transitions.

Chapter 7

Overall Discussion and Conclusion

7.1 Overall Conclusion

This thesis investigated how age, NDRT, cultural factors, and TOR design shape older drivers' ability to resume control during control transitions in Level 3 automated vehicles. Across four empirical studies, a combination of behavioural measures, self-report assessments, participatory design methods, and controlled simulator experiments were used to examine how older and younger drivers perceive automation (Chapter 3), how TOR source affects takeover performance under different NDRTs (Chapter 4), how users conceptualise TOR designs across cultures and age groups (Chapter 5), and how suspending an NDRT influences takeover performance for older drivers (Chapter 6). The findings collectively extend current understanding of control transition by demonstrating how age, NDRT cognitive demands, and TOR design features jointly shape drivers' ability to disengage from non-driving tasks, reorient attention, and resume control in Level 3 automated vehicles. These results address the Research Questions posed in Section 1.3 in the following ways:

RQ1: How do perceived trust, usefulness, and safety of Level 3 automated driving differ between younger drivers (22–25 years) and older drivers (60–80 years), and across the four older-driver age groups (60–65, 66–70, 71–75, and 76–80 years)?

Study 1 (Chapter 3) identified both similarities and differences in how younger and older drivers perceived Level 3 automation. Younger drivers viewed Level 3 vehicles positively due to their interest in novel technology and the opportunity to engage in activities such as smartphone use, social media, and email during automated driving. In contrast, older drivers primarily evaluated Level 3 automation in relation to their future mobility needs and driving capabilities. Despite these differing motivations, both age groups expressed concerns regarding trust, safety, and the ability to safely regain control following a TOR. Participants across all groups emphasised the importance of TORs that can effectively interrupt NDRTs and support a safe transition back to manual driving. These findings suggest that while younger and older drivers recognise the potential benefits of Level 3 automation, they differ substantially in their motivations for adoption and

expectations of the technology.

Study 1 demonstrated substantial heterogeneity within the older driver population. Drivers aged 60–65 generally reported confidence in their current driving abilities, enjoyed driving, and questioned the immediate value of Level 3 automation, perceiving little personal need for the technology. In contrast, drivers aged 76–80 viewed Level 3 automation more favourably, recognising it as a means of compensating for age-related declines in vision, cognition, and motor abilities and extending their driving lifespan. The intermediate age groups reflected a gradual transition between these perspectives. These findings demonstrate that attitudes toward automation, perceived usefulness, and support requirements vary considerably across later stages of ageing. Consequently, older drivers should not be treated as a homogeneous user group, and TOR design should account for the differing capabilities, motivations, and mobility needs of drivers at different stages of older adulthood.

RQ2: How does the source of TORs (IVIS vs. smartphone+IVIS) affect takeover reaction time, hazard perception, and TOR acceptance among two older driver cohorts under varying NDRT demands?

Study 2 (Chapter 4) showed that TOR source had a significant impact on takeover performance, but only under low-demand NDRTs. When the TOR was displayed both on the IVIS and on the smartphone hosting the NDRT, older drivers—particularly those aged 70+ showed faster reaction times, improved hazard perception, and higher acceptance ratings compared to when the TOR appeared only on the IVIS. However, under high-demand NDRTs, presenting the TOR on both displays did not improve performance. In these conditions, cognitive load appeared to limit the effectiveness of additional TOR redundancy. These findings indicate that the benefits of multi-source TOR presentation depend on the cognitive demands of the ongoing NDRT and that TOR design cannot rely on redundancy alone to mitigate the effects of demanding tasks.

RQ3: How do cultural (UK vs. IL) and age (young vs. old) differences impact drivers' perception of TORs and their corresponding design requirements in Level 3 automated vehicles?

Study 3 (Chapter 5) used participatory design sessions to elicit user-generated TOR concepts from younger and older drivers in the UK and Israel. The analysis revealed systematic differences in design expectations. Older drivers across both cultures preferred tri-modal TORs, explicit hazard information, and automatic interruption of NDRTs. Younger drivers emphasised gradual disruption and greater user control over the task. Cultural differences were also evident: Israeli participants consistently incorporated hazard severity indicators. These findings demonstrate that TOR expectations are shaped not only by age-related perceptual and cognitive factors but also by culturally embedded attitudes toward risk and driving. Importantly, NDRT suspension emerged as a user-proposed feature requiring empirical evaluation, motivating the final study.

RQ4: What is the impact of suspending versus continuing NDRTs during a TOR on drivers' takeover performance, hazard perception, and TOR acceptance across three age cohorts and two task demand levels?

Study 4 (Chapter 6) directly tested NDRT suspension as a system-level intervention. Suspending the NDRT at the moment of the TOR significantly improved reaction times for all age groups and for both NDRT types. However, the benefits were uneven. The most substantial gains were observed in the 70+ and 22–32 groups, while the 60–69 group showed consistently limited improvement in both reaction time and hazard perception- a non-linear age effect that challenges the treatment of older drivers as a homogeneous group and suggests that drivers in their sixties may retain sufficient cognitive flexibility to self-regulate disengagement without system intervention. Hazard perception improved only for the low-demand video NDRT; no equivalent benefit was found for the high-demand reading NDRT, indicating that cognitive load continues to constrain situational awareness even when the task is suspended. These results provide empirical evidence that automatic NDRT suspension can enhance takeover readiness but must be supplemented with additional strategies when users are engaged in cognitively demanding tasks.

7.2 Overall Contributions

This thesis makes an original and cumulative contribution to research on Take-Over Requests (TORs) in Level 3 automated vehicles by advancing empirical, methodological, and design-oriented understanding of how older drivers resume control while engaged in non-driving-related tasks (NDRTs). Positioned against existing literature, the thesis is distinctive in its explicit treatment of older drivers as a heterogeneous population, its integration of participatory design to uncover cultural differences, and its systematic examination of NDRT-aware and device-aware TOR designs.

7.2.1 Conceptual and Theoretical Contributions

This thesis advances conceptual understanding of control transition by demonstrating that chronological age alone is an insufficient basis for TOR design. Across multiple studies, older drivers were shown to differ in their motivations, attitudes toward Level 3 automation, and performance during control transitions. In particular, the findings distinguish between younger-old drivers (60-69), who exhibited greater resistance to Level 3 technology, and older-old drivers (70+), who viewed automation as a means of sustaining mobility but experienced greater re-engagement challenges. These results challenge prevailing assumptions of homogeneity within older driver cohorts and establish the need for age-sensitive and behaviourally informed models of takeover requests.

In addition, the thesis demonstrates that takeover performance is constrained by the

cognitive demands of disengaging from NDRTs, and that these constraints are shaped more strongly by task complexity than by age alone. By empirically showing that the effectiveness of TOR strategies depends on NDRT cognitive load, this work reframes NDRTs as a central moderating factor in takeover readiness rather than a secondary source of distraction.

7.2.2 Methodological Contributions

This thesis presents the first cross-cultural participatory design study of TOR interfaces conducted in a real vehicle setting, involving younger and older drivers in both the UK and Israel. Conducting design sessions inside a vehicle enabled participants to engage with spatial, embodied, and contextual aspects of TOR design, yielding requirements that are unlikely to emerge from abstract or simulator-only studies. The thesis further contributes a structured taxonomy of TOR interface components, capturing TOR messages, modalities, spatial placements, and interaction strategies across age groups, cultures, and NDRT contexts. By integrating participatory design with controlled simulator experimentation, the work establishes a design-evaluation pipeline in which user-generated concepts are empirically tested using behavioural and subjective measures. This methodological integration bridges a gap between exploratory HCI research and performance-focused human factors evaluation.

7.2.3 Empirical Contributions to TOR Design

The thesis provides novel empirical evidence regarding TOR source duality. It demonstrates that presenting the TOR simultaneously via the in-vehicle information system (IVIS) and the NDRT device significantly improves older drivers' takeover reaction times, hazard perception, usability, and TOR detectability for NDRTs with lower cognitive demand. These benefits were consistently greater for drivers aged 70+ than for those aged 60–69, indicating that attentional anchoring to the NDRT device disproportionately affects the oldest drivers.

7.2.4 System-Level Contributions: NDRT Suspension

A further original contribution of this thesis is the first empirical evaluation of suspending the NDRT at the moment of TOR issuance. The results demonstrate that NDRT suspension significantly improves takeover reaction times across age groups and task types, with the greatest benefit observed in the 70+ and 22-32 cohorts. However, the findings also show that NDRT suspension enhances hazard perception only for tasks with lower cognitive demand, such as video watching, and may increase perceived workload and reduce perceived usefulness for higher-demand tasks such as reading.

These results establish that NDRT suspension is a context-dependent intervention whose effectiveness is modulated by task complexity rather than age alone. The identification of a non-linear age effect further reinforces the need to move beyond age-based

design toward behaviour- and task-aware adaptive systems.

7.3 Limitations and Future Directions

While this thesis makes several original contributions to the understanding and design of Take-Over Requests (TORs) in Level 3 automated vehicles, a number of limitations should be acknowledged. Importantly, these limitations also highlight substantive opportunities for future research.

7.3.1 Ecological Validity and Real-World Deployment

A primary limitation of this work concerns ecological validity. Although driving simulators and in-vehicle participatory design sessions were employed, the experimental studies (Chapters 4 and 6) were conducted in controlled environments. Simulator-based evaluations cannot fully capture the complexity, unpredictability, and perceived risk associated with real-world driving, particularly in safety-critical takeover situations. Factors such as traffic density, weather conditions, road infrastructure variability, and social driving pressures may interact with TOR design and NDRT engagement in ways that are not fully reflected in simulated settings.

Future research should therefore extend the findings of this thesis to on-road or mixed-reality studies involving prototype Level 3 systems. Longitudinal field trials would be particularly valuable for examining how drivers adapt to TOR strategies over time, how repeated exposure affects trust calibration, and whether the benefits of interventions such as TOR duality or NDRT suspension persist beyond initial use.

7.3.2 Scope of NDRTs and Interaction Modalities

The empirical studies in this thesis focused on a limited set of NDRTs, primarily reading and video watching, selected to represent lower- and higher-demand visual–cognitive tasks. While these tasks are representative of common anticipated NDRTs, real-world use of Level 3 automation is likely to involve a wider and evolving range of activities, including voice-based interaction, messaging, navigation planning, gaming, and multitasking across applications.

Future work should broaden the scope of NDRTs examined and explicitly consider how different sensory, cognitive, and motor demands interact with TOR strategies. In particular, the increasing prevalence of conversational agents and auditory NDRTs raises important questions about how TORs should be coordinated across modalities and whether task suspension or TOR duality remains effective when attention is primarily auditory rather than visual.

7.3.3 Temporal Dynamics and Learning Effects

Another limitation of the present work is the absence of longitudinal assessment. The studies examined driver behaviour and perception during single-session interactions with Level 3 automation. As a result, the findings do not capture how learning, familiarity, or changes in cognitive strategy over time may influence takeover performance. For example, drivers may learn to anticipate TORs, adjust their engagement with NDRTs, or develop trust calibration strategies that alter the effectiveness of specific TOR designs.

Future research should investigate the temporal dynamics of takeover behaviour through repeated exposure and long-term use. Such work would allow examination of whether age-related differences diminish or intensify with experience, and whether adaptive TOR systems should evolve alongside the user rather than remaining static.

7.3.4 Cultural Generalisability

The cross-cultural participatory design study presented in Chapter 5 compared drivers from the UK and Israel, revealing meaningful differences in TOR expectations, urgency framing, and risk communication. While this comparison provides valuable insights, it represents only a subset of global driving cultures. Driving norms, regulatory frameworks, and attitudes toward automation vary substantially across regions, particularly between Western and non-Western contexts.

Future research should extend this work to a broader range of cultural settings, including regions with different enforcement regimes, traffic norms, and levels of automation acceptance. Expanding cultural coverage would support the development of TOR systems that balance global consistency with local adaptation and would clarify which design principles generalise across cultures and which require regional customisation.

7.3.5 Age as a Behavioural Rather Than Chronological Variable

Although this thesis demonstrated that older drivers should not be treated as a homogeneous group, age was still operationalised using chronological categories. The identification of non-linear age effects—particularly the finding that the 70+ group benefited more consistently from certain interventions than the 60–69 group—suggests that chronological age may be a coarse proxy for underlying cognitive and behavioural characteristics.

Future research should explore behaviour-based and capability-based models of takeover readiness that incorporate cognitive flexibility, task-switching ability, and attentional control rather than age alone. Integrating such models with real-time driver state monitoring could enable TOR systems that adapt dynamically to individual drivers rather than predefined age groups.

7.3.6 Regulatory and Ethical Considerations

Finally, while the thesis discusses regulatory and practical constraints surrounding interventions such as NDRT suspension, these aspects were not empirically evaluated. Au-

tomatically suspending user activities raises important questions related to user autonomy, consent, and liability, particularly in commercial deployment. Future work should therefore engage with regulatory bodies, industry stakeholders, and users to explore acceptable boundaries for system intervention and to evaluate how such features can be implemented transparently and ethically.

7.3.7 Summary

In summary, while the present thesis provides robust empirical and design-oriented contributions to the understanding of TORs in Level 3 automated vehicles, it also opens multiple avenues for future research. By highlighting the roles of task demand, age heterogeneity, cultural context, and system-level intervention, this work establishes a foundation for the next generation of adaptive, user-aware takeover support systems. Addressing the limitations identified here will be essential for translating experimental findings into real-world Level 3 deployments that are safe, inclusive, and responsive to the diversity of human drivers.

Appendices

Appendix A1

Introduction

Thank you for joining me today. We have come together to discuss your attitudes, thoughts, and perceptions of Level 3 autonomous cars. We will also talk about the types of non-driving related tasks you would be willing to engage in while travelling in a Level 3 vehicle.

I am interested in hearing how Level 3 vehicles might help improve your travel experience, as well as your ideas, concerns, fears, and questions about these cars.

Before we begin, does anyone have any questions? (Please address any questions before moving on.)

I will read through several questions. Please feel free to respond openly and offer any feedback or additional information as we go along.

Background Questions

1. How often do you use a car?
2. What mode of transportation do you regularly use for your trips or errands?
3. Do you prefer to drive or to be a passenger?

Definition of a Level 3 Autonomous Car

In Level 3 autonomous cars, the automated driving system is capable of and responsible for driving, but it relies on the human driver to respond appropriately to a *Take Over Request* under certain road conditions. The Take Over Request will be communicated clearly by the vehicle.

In Level 3 automation, the human driver does not need to constantly monitor the road and is free to engage in non-driving related tasks (such as reading or using a mobile phone). However, the driver is still required to remain available and intervene when the system reaches its operational limits.

Non-Driving Related Tasks (Baseline)

1. How often do you send emails in your daily life? If not currently, did you do this in the past?
2. Do you use social media? If yes, which platforms do you use?
3. Do you read, knit, or talk on the phone? What kinds of social or leisure activities do you generally prefer to engage in?

Scenario 1

The driver enters the address of a friend's new apartment into the navigation system. To reach the motorway, the car is manually controlled by the driver. The car is then switched to autonomous mode and handles accelerating, steering, and braking completely autonomously. The driver can take their hands off the steering wheel and watch a film in a relaxed manner.

An hour later, before reaching the motorway exit, the car reminds the driver to take back control. The driver takes control of the driving and arrives safely at the destination.

1. How do you think you would feel driving a Level 3 autonomous car in this scenario?
2. How do you think the Take Over Request should be communicated to you?
3. What non-driving related tasks would you choose to do in this situation?

Scenario 2

Lauren drives her son to school daily. She selects the school address from the navigation system. The car enters autonomous mode and starts driving toward the school. In the meantime, Lauren is reading and signing a consent form for her son to join a school trip.

Ten minutes later, when approaching the school vicinity, the car issues a Take Over Request.

1. How do you think you would feel driving a Level 3 autonomous car in this scenario?
2. How do you think the Take Over Request should be communicated to you?
3. What non-driving related tasks would you do in this situation?

Interaction Design

1. From an interaction design perspective, what would encourage you to use a Level 3 autonomous car?
2. From an interaction design perspective, what would encourage you to engage in non-driving related tasks while travelling in a Level 3 car?

Interface Examples

Participants will be shown images of interior interfaces of autonomous vehicles. These examples will be used to support discussion and help explain interaction design concepts collaboratively with participants.

General Questions

1. What excites you about Level 3 autonomous cars?
    - What positive impacts do you think Level 3 cars may have on your life or on the lives of your family members?
    - Do you think they could improve your social life?
    - Do you think they could increase your free time?
2. What concerns or fears do you have about using or owning a Level 3 autonomous car?
    - Do you worry about accidents and who would be at fault?
    - Are you concerned about your personal safety?
    - Are you worried about being able to communicate effectively with the car?
3. If you were to decide to use autonomous vehicles, how often do you think you would do so?
    - Can you give examples of trips where you would choose to use these vehicles?
4. If you had to choose the single most important aspect that Level 3 autonomous vehicles should incorporate into their design, what would it be?
5. What specific design features would increase your willingness to use a Level 3 car?
6. Would learning more about Level 3 capabilities and practicing driving a car like this increase your willingness to use it?

Appendix A2

Table 7.1: Descriptive statistics and pairwise comparisons between older age groups for perceived trust, safety, and usefulness.

Measure	Age group	Mean	SE	Pairwise comparisons (p)
Trust	60–65	2.87	0.112	60–65 vs. 66–70: .999
	66–70	2.92	0.112	60–65 vs. 71–75: 1.000
	71–75	2.86	0.111	60–65 vs. 76–80: .819
	76–80	3.05	0.114	66–70 vs. 71–75: .996
				66–70 vs. 76–80: .927
				71–75 vs. 76–80: .765
Safety	60–65	2.49	0.145	60–65 vs. 66–70: .890
	66–70	2.68	0.145	60–65 vs. 71–75: .996
	71–75	2.56	0.143	60–65 vs. 76–80: .464
	76–80	2.83	0.146	66–70 vs. 71–75: .982
				66–70 vs. 76–80: .947
				71–75 vs. 76–80: .698
Usefulness	60–65	2.47	0.134	60–65 vs. 66–70: .928
	66–70	2.62	0.134	60–65 vs. 71–75: .999
	71–75	2.52	0.133	60–65 vs. 76–80: .744
	76–80	2.70	0.135	66–70 vs. 71–75: .982
				66–70 vs. 76–80: .994
				71–75 vs. 76–80: .874

Appendix B

Advanced transport telematics acceptance assessment

Table 1. The nine items in the questionnaire

My judgements of the (...) system are... (please tick a box on every line)		
1	useful	useless
2	pleasant	unpleasant
3	bad	good
4	nice	annoying
5	effective	superfluous
6	irritating	likeable
7	assisting	worthless
8	undesirable	desirable
9	raising alertness	sleep-inducing

Figure 7.1: User Acceptance of New Transport Telematics

Bibliography

[1] Sae j3016tm levels of driving automation, 2021.

[2] Hillary Abraham, Chaiwoo Lee, Samantha Brady, Craig Fitzgerald, Bruce Mehler, Bryan Reimer, and Joseph Coughlin. Autonomous vehicles and alternatives to driving: Trust, preferences, and effects of age. 01 2017.

[3] Ammar Al-Taie, Graham Wilson, Frank Pollick, and Stephen Anthony Brewster. Pimp my ride: Designing versatile ehmis for cyclists. In *Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '23, page 213–223, New York, NY, USA, 2023. Association for Computing Machinery.

[4] Ibrahim Alsghan, Uneb Gazder, Khaled Assi, Gazi Hassan Hakem, Mohammad Abduljalil Sulail, and Osamah Abdulrahman Alsuhaibani. The determinants of consumer acceptance of autonomous vehicles: A case study in riyadh, saudi arabia. *International Journal of Human-Computer Interaction*, 38:1375–1387, 2022.

[5] Rahaf Alsnih and David A. Hensher. The mobility and accessibility expectations of seniors in an aging population. *Transportation Research Part A: Policy and Practice*, 37:903–916, 2003.

[6] Shruti Amre and Ye Sun. Effects of non-driving task related workload and situational awareness in semi-autonomous vehicles. In Julia L. Wright, Daniel Barber, Sofia Scataglini, and Sudhakar L. Rajulu, editors, *Advances in Simulation and Digital Human Modeling*, pages 247–254, Cham, 2021. Springer International Publishing.

[7] Kaarin J. Anstey, Joanne Wood, Stephen Lord, and Janine G. Walker. Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Review*, 25(1):45–65, 2005.

[8] Bima Pramudya Asaddulloh, Afrig Aminuddin, Majid Rahardi, Ferian Fauzi Abdulloh, Dhanar Intan Surya Saputra, and Faza Ridha Alfana. Ui/ux development based on design thinking framework for mentor on demand service. In *2023 6th International Conference on Information and Communications Technology (ICOIACT)*, pages 120–125, 2023.

[9] Kamil Bartoń. *MuMIn: Multi-Model Inference*, 2023. R package version 1.50.1.

[10] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

[11] Patrick J. Bennett, Robert Sekuler, and Allison B. Sekuler. The effects of aging on motion detection and direction identification. *Vision Research*, 47(6):799–809, 2007.

[12] Amitai Y Bin-Nun, Patricia Derler, Noushin Mehdipour, and Radboud Duintjer Tebbens. How should autonomous vehicles drive? policy, methodological, and social considerations for designing a driver. *Humanities and social sciences communications*, 9(1):1–13, 2022.

[13] David Bissell, Thomas Birtchnell, Anthony Elliott, and Eric L. Hsu. Autonomous automobilities: The social impacts of driverless vehicles. *Current Sociology*, 68:116–134, 1 2020.

[14] James P. Bliss and Mariea C. Dunn. Behavioural implications of alarm mistrust as a function of task workload. *Ergonomics*, 43(9):1283–1300, 2000. PMID: 11014752.

[15] Shadan Sadeghian Borojeni, Lars Weber, Wilko Heuten, and Susanne Boll. From reading to driving: Priming mobile users for take-over situations in highly automated driving. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '18, New York, NY, USA, 2018. Association for Computing Machinery.

[16] Avinoam Borowsky, William J. Horrey, Yulan Liang, Angela Garabet, Lucinda Simmons, and Donald L. Fisher. The effects of brief visual interruption tasks on drivers' ability to resume their visual search for a pre-cued hazard. *Accident Analysis and Prevention*, 93:207–216, 8 2016.

[17] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.

[18] L. Mark Carrier, Larry D. Rosen, Nancy A. Cheever, and Alex F. Lim. Causes, effects, and practicalities of everyday multitasking. *Developmental Review*, 35:64–78, 2015. Special Issue: Living in the "Net" Generation: Multitasking, Learning, and Development.

[19] GB Christopher. The licensing and safety of older drivers in britain. *Accident Analysis and Prevention*, 50:732–741, 2013.

[20] Hallie Clark and Jing Feng. Age differences in the takeover of vehicle control and engagement in non-driving-related activities in simulated driving with conditional automation. *Accident Analysis and Prevention*, 106:468–479, 9 2017.

[21] Sherrilene Classen, Justin Mason, Seung Woo Hwangbo, James Wersal, Jason Rogers, and Virginia Sisiopiku. Older drivers' experience with automated vehicle technology. *Journal of Transport and Health*, 22, 9 2021.

[22] Mitchell L. Cunningham, Michael A. Regan, Timothy Horberry, K. Weeratunga, and Vinayak Dixit. Public opinion about automated vehicles in australia: Results from a large-scale national survey. *Transportation Research Part A: Policy and Practice*, 129:1–18, 11 2019.

[23] Sara J. Czaja, Walter R. Boot, Neil Charness, and Wendy A. Rogers. *Designing for Older Adults: Principles and Creative Human Factors Approaches*. CRC Press, Boca Raton, FL, 3 edition, 2019.

[24] Frank N. Dempster. The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, 12:45–75, 1992.

[25] Rosa DeRamus. *The effect of driver age and experience on risk assessment and risk prediction*. University of Massachusetts Amherst, 2005.

[26] Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. Taming the ehmi jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives*, 7:100174, 2020.

[27] Yongqi Dong, Chang Liu, Yiyun Wang, and Zhe Fu. Towards understanding worldwide cross-cultural differences in implicit driving cues: Review, comparative analysis, and research roadmap, 2024.

[28] Driving and Vehicle Standards Agency. Theory test: cars, 2023.

[29] Mike Duncan, Neil Charness, Tim Chapin, Mark Horner, Lindsay Stevens, Amanda Richard, Dustin Souders, Jeremy Crute, Alex Riemondy, Donald Morgan, and et al. Enhanced mobility for aging populations using automated vehicles., Dec 2015.

[30] Mahdi Ebnali, Pedram Ahmadnezhad, Alireza Shateri, Adel Mazloumi, Majid Ebnali Heidari, and Ahmad Reza Nazeri. The effects of cognitively demanding dual-task driving condition on elderly people's driving performance; real driving monitoring. *Accident Analysis and Prevention*, 94:198–206, 2016.

[31] Mahdi Ebnali, Pedram Ahmadnezhad, Alireza Shateri, Adel Mazloumi, Majid Ebnali Heidari, and Ahmad Reza Nazeri. The effects of cognitively demanding dual-task driving condition on elderly people's driving performance; real driving monitoring. *Accident Analysis and Prevention*, 94:198–206, 2016.

[32] Fredrick Ekman, Mikael Johansson, and Jana Sochor. Creating appropriate trust in automated vehicle systems: A framework for hmi design. *IEEE Transactions on Human-Machine Systems*, 48:95–101, 2 2018.

[33] Mica R Endsley. Measurement of situation awareness in dynamic systems. *Human factors*, 37(1):65–84, 1995.

[34] Daniel J. Fagnant and Kara Kockelman. Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77:167–181, 7 2015.

[35] S Fairclough. Psychophysiological measures of workload. *Driving future vehicles*, page 362, 1993.

[36] L. Ferreira, R. Ferreira Santos-Galduróz, C. P. Ferri, and J. C. Fernandes Galduróz. Rate of cognitive decline in relation to sex after 60 years-of-age: A systematic review. *Geriatrics & Gerontology International*, 14(1):23–31, 2014.

[37] Michela Folli and Joanne M. Bennett. Factors that predict hazard perception in older adult road-users: A systematic review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 99:19–35, 2023.

[38] Hannah J. Foy and Peter Chapman. Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation. *Applied Ergonomics*, 73:90–99, 2018.

[39] Anna-Katharina Frison, Philipp Wintersberger, and Andreas Riener. First person trolley problem: Evaluation of drivers' ethical decisions in a driving simulator. In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '16 Adjunct, page 117–122, New York, NY, USA, 2016. Association for Computing Machinery.

[40] C. Gasne, L. Paire-Ficout, S. Bordel, S. Lafont, and M. Ranchet. Takeover performance of older drivers in automated driving: A review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 87:347–364, 2022.

[41] Pnina Gershon, Kellienne R. Sita, Chunming Zhu, Johnathon P. Ehsani, Sheila G. Klauer, Tom A. Dingus, and Bruce G. Simons-Morton. Distracted driving, visual inattention, and crash risk among teenage drivers. *American Journal of Preventive Medicine*, 56:494–500, 4 2019.

[42] Christian Gold, Daniel Damböck, Klaus Bengler, and Lutz Lorenz. Partially automated driving as a fallback level of high automation. In *6. tagung fahrerassistenzsysteme*, 2013.

[43] Christian Gold, Moritz Körber, Christoph Hohenberger, David Lechner, and Klaus Bengler. Trust in automation – before and after the experience of take-over scenarios in a highly automated vehicle. *Procedia Manufacturing*, 3:3025–3032, 2015.

[44] Christian Gold, Moritz Körber, David Lechner, and Klaus Bengler. Taking over control from highly automated vehicles in complex traffic situations: The role of traffic density. *Human Factors*, 58(4):642–652, 2016. PMID: 26984515.

[45] Noah J. Goodall. *Machine Ethics and Automated Vehicles*, pages 93–102. Springer International Publishing, Cham, 2014.

[46] Thomas Goodge, Jingyi Li, Shiv G Patel, Victor Gravis, and Soyeon Kim. Enhancing explainability of automated vehicles with culturally-aware hmis. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '23 Adjunct, page 337–339, New York, NY, USA, 2023. Association for Computing Machinery.

[47] Thomas Alexander Goodge, Frank Pollick, and Stephen Anthony Brewster. Can you hazard a guess?: Evaluating the effect of augmented reality cues on driver hazard prediction. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, 2024. Association for Computing Machinery.

[48] Sean E Goodison, Jeremy D Barnum, Michael JD Vermeer, Dulani Woods, Tatiana Lloyd-Dotta, and Brian A Jackson. *Autonomous Road Vehicles and Law Enforcement*. RAND Corporation, 2020.

[49] UK Goverment. Statistical data set: Reported road collisions, vehicles and casualties tables for great britain, December 2022.

[50] Iwona Grobelna, David Mailland, and Mikołaj Horwat. Design of automotive hmi: New challenges in enhancing user experience, safety, and security. *Applied Sciences*, 15(10):5572, 2025.

[51] Danan Gu, Kirill Andreev, and Matthew E. Dupre. Major trends in population growth around the world. *China CDC Weekly*, 3(27):604, 2021.

[52] Siti Haerani, Rika Dwi Ayu Parmitasari, Elsina Huberta Aponno, and Zany Irayati Aunalal. Moderating effects of age on personality, driving behavior towards driving outcomes. *International Journal of Human Rights in Healthcare*, 12(2):91–104, 2019.

[53] Michael Harley. Testing (and trusting) mercedes-benz level 3 drive pilot in germany, Aug 2022.

[54] Sandra G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006.

[55] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[56] Sebastian Hergeth, Lutz Lorenz, Josef F Krems, and Lars Toenert. Effects of takeover requests and cultural background on automation trust in highly automated driving. In *Driving assessment conference*, volume 8. University of Iowa, 2015.

[57] Kai Holländer and Bastian Pfleging. Preparing drivers for planned control transitions in automated cars. In *Proceedings of the 17th international conference on mobile and ubiquitous multimedia*, pages 83–92, 2018.

[58] George Holley-Moore and Helen Creighton. The future of transport in an ageing society. Report, International Longevity Centre UK and Age UK, 2015. Accessed: 2025-01-13.

[59] William Horrey and Christopher Wickens. Multiple resource modeling of task interference in vehicle control, hazard awareness and in-vehicle task performance. 01 2003.

[60] Mark S. Horswill, Shelby A. Marrington, Cynthia M. McCullough, Joanne Wood, Nancy A. Pachana, Jenna McWilliam, and Maria K. Raikos. The Hazard Perception Ability of Older Drivers. *The Journals of Gerontology: Series B*, 63(4):P212–P218, 07 2008.

[61] MARK S. HORSWILL, NANCY A. PACHANA, JOANNE WOOD, SHELBY A. MARRINGTON, JENNA McWILLIAM, and CYNTHIA M. McCULLOUGH. A comparison of the hazard perception ability of matched groups of healthy drivers aged 35 to 55, 65 to 74, and 75 to 84 years. *Journal of the International Neuropsychological Society*, 15(5):799–802, 2009.

[62] Gaojian Huang, Ya-Hsin Hung, Robert W. Proctor, and Brandon J. Pitts. Age is more than just a number: The relationship among age, non-chronological age factors, self-perceived driving abilities, and autonomous vehicle acceptance. *Accident Analysis and Prevention*, 178:106850, 12 2022.

[63] Gaojian Huang, Maya Luster, Ilayda Karagol, Jun Woo Park, and Brandon J. Pitts. Self-perception of driving abilities in older age: A systematic review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 74:307–321, 10 2020.

[64] Gaojian Huang and Brandon Pitts. Age-related differences in takeover request modality preferences and attention allocation during semi-autonomous driving. In Qin Gao and Jia Zhou, editors, *Human Aspects of IT for the Aged Population. Technologies, Design and User Experience*, pages 135–146, Cham, 2020. Springer International Publishing.

[65] Gaojian Huang and Brandon J. Pitts. The effects of age and physical exercise on multimodal signal responses: Implications for semi-autonomous vehicle takeover requests. *Applied Ergonomics*, 98:103595, 2022.

[66] Gaojian Huang and Brandon J. Pitts. To inform or to instruct? an evaluation of meaningful vibrotactile patterns to support automated vehicle takeover performance. *IEEE Transactions on Human-Machine Systems*, 53(4):678–687, 2023.

[67] Gaojian Huang, Clayton Steele, Xinrui Zhang, and Brandon J. Pitts. Multimodal cue combinations: A possible approach to designing in-vehicle takeover requests for semi-autonomous driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63:1739–1743, 11 2019.

[68] Carrie Huisingh, Emily B. Levitan, Marguerite R. Irvin, Paul MacLennan, Virginia Wadley, and Cynthia Owsley. Visual sensory and visual-cognitive function and rate of crash and near-crash involvement among older drivers using naturalistic driving data. *Investigative Ophthalmology Visual Science*, 58(7):2959–2967, 06 2017.

[69] Yeji Hwang and Gwi Ryung Son Hong. Predictors of driving cessation in community-dwelling older adults: A 3-year longitudinal study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 52:202–209, 1 2018.

[70] Christian P. Janssen, Shamsi T. Iqbal, Andrew L. Kun, and Stella F. Donker. Interrupted by my car? implications of interruption and interleaving research for automated vehicles. *International Journal of Human-Computer Studies*, 130:221–233, 2019.

[71] Christian P. Janssen, Shamsi T. Iqbal, Andrew L. Kun, and Stella F. Donker. Interrupted by my car? implications of interruption and interleaving research for automated vehicles. *International Journal of Human-Computer Studies*, 130:221–233, 2019.

[72] Marie Jaussein, Lucie Lévêque, Jonathan Deniel, Thierry Bellet, Hélène Tattegrain, and Claude Marin-Lamellet. How do non-driving-related tasks affect engagement under highly automated driving situations? a literature review. *Frontiers in Future Transportation*, 2, 07 2021.

[73] Marie Jaussein, Lucie Lévêque, Jonathan Deniel, Thierry Bellet, Hélène Tattegrain, and Claude Marin-Lamellet. How do non-driving-related tasks affect engagement under highly automated driving situations? a literature review. *Frontiers in Future Transportation*, Volume 2 - 2021, 2021.

[74] Alvin Jude. How to improve the self-driving car passenger experience - ericsson, Jan 2020.

[75] Keita Kamijo and Yuji Takeda. Regular physical activity improves executive function during task switching in young adults. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, 75(3):304—311, March 2010.

[76] Melanie Karthaus and Michael Falkenstein. Functional changes and driving performance in older drivers: Assessment and interventions, 6 2016.

[77] Melanie Karthaus and Michael Falkenstein. Functional changes and driving performance in older drivers: Assessment and interventions. *Geriatrics*, 1(2), 2016.

[78] Kanwaldeep Kaur and Giselle Rampersad. Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. *Journal of Engineering and Technology Management - JET-M*, 48:87–96, 4 2018.

[79] Fareed Kaviani, Kristie L. Young, Brady Robards, and Sjaan Koppel. "like it's wrong, but it's not that wrong:" exploring the normalization of risk-compensatory strategies among young drivers engaging in illegal smartphone use. *Journal of Safety Research*, 78:292–302, 9 2021.

[80] Bilal Alam Khan, Maria Chiara Leva, and Sam Cromie. A systematic review of older drivers in a level 3 autonomous vehicle: A cognitive load perspective. In Luca Longo and Maria Chiara Leva, editors, *Human Mental Workload: Models and Applications*, pages 60–77, Cham, 2021. Springer International Publishing.

[81] Andrea Kiesel, Marco Steinhauser, Mike Wendt, Michael Falkenstein, Kerstin Jost, Andrea M Philipp, and Iring Koch. Control and interference in task switching—a review. *Psychological bulletin*, 136(5):849, 2010.

[82] Soyeon Kim, René van Egmond, and Riender Happee. Effects of user interfaces on take-over performance: A review of the empirical evidence. *Information*, 12(4), 2021.

[83] Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26, 2017.

[84] M. Kyriakidis, R. Happee, and J. C.F. De Winter. Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, 32:127–140, 6 2015.

[85] Jan Luca Kästle, Bani Anvari, Jan Peters, Jakub Krol, and Helge Wurdemann. Correlation between reaction time, multi-modal feedback and take-over requests for level 3 automated vehicles. *IEEE Transactions on Haptics*, 18(3):776–782, 2025.

[86] M. König and L. Neumayr. Users' resistance towards radical innovations: The case of the self-driving car. *Transportation Research Part F: Traffic Psychology and Behaviour*, 44:42–52, 1 2017.

[87] Moritz Körber, Christian Gold, David Lechner, and Klaus Bengler. 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, 2016.

[88] Patrick Langdon, Ioannis Politis, Mike Bradley, Lee Skrypchuk, Alex Mouzakitis, and John Clarkson. Obtaining design requirements from the public understanding of driverless technology. volume 597, pages 749–759. Springer Verlag, 2018.

[89] Seul Chan Lee, Sol Hee Yoon, and Yong Gu Ji. Effects of non-driving-related task attributes on takeover quality in automated vehicles. *International Journal of Human–Computer Interaction*, 37(3):211–219, 2021.

[90] Guohua Li, Elisa Braver, and Li-Hui Chen. Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. *Accident; analysis and prevention*, 35:227–35, 04 2003.

[91] Shuo Li, Phil Blythe, Weihong Guo, and Anil Namdeo. Investigating the effects of age and disengagement in driving on driver's takeover control performance in highly automated vehicles. *Transportation planning and technology*, 42(5):470–497, 2019.

[92] Shuo Li, Phil Blythe, Weihong Guo, and Anil Namdeo. 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, 4 2019.

[93] Shuo Li, Phil Blythe, Yanghanzi Zhang, Simon Edwards, Jin Xing, Weihong Guo, Yanjie Ji, Paul Goodman, and Anil Namdeo. 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, 4 2021.

[94] Xian Liu, Ruth Madigan, Ehsan Sadraei, Yee Mun Lee, and Natasha Merat. Drivers' engagement in ndrts during automated driving linked to travelling speed and surrounding traffic. *Transportation Research Part F: Traffic Psychology and Behaviour*, 101:332–339, 2024.

[95] Stephen Lyman, Susan A Ferguson, Elisa R Braver, and Allan F Williams. Older driver involvements in police reported crashes and fatal crashes: trends and projections. *Injury prevention*, 8(2):116–120, 2002.

[96] Davide Maggi, Richard Romano, Oliver Carsten, and Joost de Winter. When terminology hinders research: the colloquialisms of transitions of control in automated driving. *Cognition, Technology Work*, 24:1–12, 08 2022.

[97] Dominique Makowski, Daniel Lüdecke, Indrajeet Patil, Rémi Thériault, Mattan S. Ben-Shachar, and Brenton M. Wiernik. Automated results reporting as a practical tool to improve reproducibility and methodological best practices adoption. *CRAN*, 2023.

[98] Sampada Marathe and S. Shyam Sundar. What drives customization? control or identity? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 781–790, New York, NY, USA, 2011. Association for Computing Machinery.

[99] Richard A. Marottoli, Carlos F. Mendes De Leon, Thomas A. Glass, Christianna S. Williams, Leo M. Cooney, Lisa F. Berkman, and Mary E. Tinetti. Driving cessation and increased depressive symptoms: Prospective evidence from the new haven epese. *Journal of the American Geriatrics Society*, 45:202–206, 1997.

[100] Marsh & McLennan Companies, Asia Pacific Risk Center. Advancing into the golden years: Cost of healthcare for asia pacific's elderly, 2016. [Accessed: 2021-06-16].

[101] Daniel McFadden. Quantitative methods for analysing travel behaviour of individuals: some recent developments. In *Behavioural travel modelling*, pages 279–318. Routledge, 2021.

[102] Gerald McGwin, Jr and David B. Brown. Characteristics of traffic crashes among young, middle-aged, and older drivers. *Accident Analysis and Prevention*, 31(3):181–198, 1999.

[103] Vivien Melcher, Stefan Rauh, Frederik Diederichs, Harald Widlroither, and Wilhelm Bauer. Take-over requests for automated driving. *Procedia Manufacturing*, 3:2867–2873, 2015.

[104] Natasha Merat, A. Hamish Jamson, Frank C.H. Lai, and Oliver Carsten. Highly automated driving, secondary task performance, and driver state. *Human Factors*, 54:762–771, 10 2012.

[105] Natasha Merat, A Hamish Jamson, Frank CH Lai, Michael Daly, and Oliver MJ Carsten. Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation research part F: traffic psychology and behaviour*, 27:274–282, 2014.

[106] Akira Miyake, Naomi P. Friedman, Michael J. Emerson, Alexander H. Witzki, Amy Howerter, and Tor D. Wager. 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, 2000.

[107] Nichole Morris, Curtis Craig, David Libby, and Jennifer Cooper. Older driver support system (odss) usability and design investigation, 01 2018.

[108] Andreas Müller, Natacha Fernandes-Estrela, Ruben Hetfleisch, Lukas Zecha, and Bettina Abendroth. Effects of non-driving related tasks on mental workload and take-over times during conditional automated driving. *European Transport Research Review*, 13:16, 12 2021.

[109] Ziad Nasreddine, Natalie Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey Cummings, and Howard Chertkow. The montreal cognitive assessment, moca: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53:695–9, 05 2005.

[110] Sina Nordhoff, Jork Stapel, Xiaolin He, Alexandre Gentner, and Riender Happee. Perceived safety and trust in sae level 2 partially automated cars: Results from an online questionnaire. *PLoS ONE*, 16, 12 2021.

[111] Erfan Pakdamanian, Erzhen Hu, Shili Sheng, Sarit Kraus, Seongkook Heo, and Lu Feng. Enjoy the ride consciously with cawa: Context-aware advisory warnings for automated driving. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '22, page 75–85, New York, NY, USA, 2022. Association for Computing Machinery.

[112] Denise C. Park and Patricia Reuter-Lorenz. The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60(1):173–196, 2009. PMID: 19035823.

[113] UK Parliament. Autonomous vehicle bill 2023., 2023.

[114] William Payre, Julien Cestac, and Patricia Delhomme. Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27:252–263, 2014.

[115] Jonathan W Peirce. Psychopy—psychophysics software in python. *Journal of neuroscience methods*, 162(1-2):8–13, 2007.

[116] Qijia Peng and Sunao Iwaki. Visual attention of young and older drivers in takeover tasks of highly automated driving. In *Human Aspects of IT for the Aged Population. Technologies, Design and User Experience: 6th International Conference, ITAP 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I 22*, pages 210–221. Springer, 2020.

[117] Qijia Peng, Yanbin Wu, Nan Qie, and Sunao Iwaki. Age-related effects of executive function on takeover performance: A simulator study in automated driving. 2021.

[118] Bastian Pfleging, Maurice Rang, and Nora Broy. Investigating user needs for non-driving-related activities during automated driving. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, MUM '16, page 91–99, New York, NY, USA, 2016. Association for Computing Machinery.

[119] Brandon J. Pitts and Nadine Sarter. What you don't notice can harm you: Age-related differences in detecting concurrent visual, auditory, and tactile cues. *Human Factors*, 60(4):445–464, 2018. PMID: 29470102.

[120] Ioannis Politis, Stephen Brewster, and Frank Pollick. Evaluating multimodal driver displays under varying situational urgency. pages 4067–4076. Association for Computing Machinery, 2014.

[121] Ioannis Politis, Stephen Brewster, and Frank Pollick. Language-based multimodal displays for the handover of control in autonomous cars. In *Proceedings of the 7th*

International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI '15, page 3–10, New York, NY, USA, 2015. Association for Computing Machinery.

[122] Ioannis Politis, Stephen Brewster, and Frank Pollick. Using multimodal displays to signify critical handovers of control to distracted autonomous car drivers. *Int. J. Mob. Hum. Comput. Interact.*, 9(3):1–16, jul 2017.

[123] Alexander Pollatsek, Matthew R. E. Romoser, and Donald L. Fisher. Identifying and remediating failures of selective attention in older drivers. *Current Directions in Psychological Science*, 21(1):3–7, 2012. PMID: 23082045.

[124] Jonas Radlmayr, Christian Gold, Lutz Lorenz, Mehdi Farid, and Klaus Bengler. 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, 2014.

[125] Andry Rakotonirainy, Dale Steinhardt, Patricia Delhomme, Millie Darvell, and Amy Schramm. Older drivers' crashes in queensland, australia. *Accident Analysis and Prevention*, 48:423–429, 2012.

[126] Jan Reich and Mario Trapp. Sinadra: Towards a framework for assurable situation-aware dynamic risk assessment of autonomous vehicles. In *2020 16th European Dependable Computing Conference (EDCC)*, pages 47–50, Sep. 2020.

[127] Robyn D. Robertson, Heather Woods-Fry, Ward G.M. Vanlaar, and Marisela Mainegra Hing. Automated vehicles and older drivers in canada. *Journal of Safety Research*, 70:193–199, 9 2019.

[128] Luka Rukonić, Marie-Anne Pungu Mwange, and Suzanne Kieffer. How older drivers perceive warning alerts? insights for the design of driver–car interaction. *SN Comput. Sci.*, 4(1), nov 2022.

[129] Dario D. Salvucci, Niels A. Taatgen, and Jelmer P. Borst. Toward a unified theory of the multitasking continuum: from concurrent performance to task switching, interruption, and resumption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, page 1819–1828, New York, NY, USA, 2009. Association for Computing Machinery.

[130] Holger Schielzeth, Niels J Dingemanse, Shinichi Nakagawa, David F Westneat, Hassen Allegue, Céline Teplitsky, Denis Réale, Ned A Dochtermann, László Zsolt Garamszegi, and Yimen G Araya-Ajoy. Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in ecology and evolution*, 11(9):1141–1152, 2020.

[131] Bernhard Schlag, Sebastian Werth, Benedikt Gahr, Rene Meir, Sandra Trösterer, and Simon Kopp. Review of studies on older drivers' behavior and stress—methods, results, and outlook. *Sensors*, 21(10):3503, 2021.

[132] Christina Schneegass and Fiona Draxler. *Designing Task Resumption Cues for Interruptions in Mobile Learning Scenarios*, pages 125–181. Springer International Publishing, Cham, 2021.

[133] Brandon Schoettle and Michael Sivak. A survey of public opinion about autonomous and self-driving vehicles in the U.S., the U.K., and australia. Technical Report UMTRI-2014-21, The University of Michigan Transportation Research Institute, Ann Arbor, MI, USA, 2014. Technical Report.

[134] Emre Sezgin and Simon Lin. Technology-based interventions, assessments, and solutions for safe driving training for adolescents: Rapid review. *JMIR Mhealth Uhealth*, 7(1):e11942, Jan 2019.

[135] Lipaz Shamoa-Nir and Uri Durchin. Israeli driving culture: Driving style and drivers' attitudes towards traffic laws. *Israeli Sociology*, 17(2):138–159, 2016.

[136] Abbas Sheykhfard, Farshidreza Haghighi, Grigorios Fountas, Subasish Das, and Ali Khanpour. How do driving behavior and attitudes toward road safety vary between developed and developing countries? evidence from iran and the netherlands. *Journal of Safety Research*, 85:210–221, 2023.

[137] Elisabeth Shi and Klaus Bengler. 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, 2022.

[138] Charles Singletary. Panasonic's autonomous car design includes an ar equipped cockpit, Mar 2023.

[139] Vera Skvirsky, Orit Taubman Ben-Ari, Timothy J. Greenbury, and Carlo G. Prato. Contributors to young drivers' driving styles – a comparison between israel and queensland. *Accident Analysis and Prevention*, 109:47–54, 2017.

[140] Rawan Srour Zreik, Thomas Goodge, Monika Harvey, and Stephen Brewster. Where's my tor?: Evaluating the effect of take-over request source on older drivers' control transition in level 3 cars. *Proc. ACM Hum.-Comput. Interact.*, 8(MHCI), September 2024.

[141] Rawan Srour Zreik, Monika Harvey, and Stephen Anthony Brewster. Age matters: Investigating older drivers' perception of level 3 autonomous cars as a heterogeneous age group. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '23

Adjunct, page 13–18, New York, NY, USA, 2023. Association for Computing Machinery.

[142] Rawan Srour Zreik, Monika Harvey, and Stephen Anthony Brewster. All-inclusive tors: Cross-cultural and age-sensitive design for take-over requests in level 3 cars. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA, 2025. Association for Computing Machinery.

[143] Yaakov Stern. Cognitive reserve. *Neuropsychologia*, 47(10):2015–2028, 2009.

[144] Gregor Strle, Yilun Xing, Erika E. Miller, Linda Ng Boyle, and Jaka Sodnik. Take-over time: A cross-cultural study of take-over responses in highly automated driving. *Applied Sciences*, 11(17), 2021.

[145] Chinebuli Uzondu, Samantha Jamson, Daryl Hibberd, and AECOM Strategic Consultancy. Cross-cultural differences in drivers' behaviour: a driving simulator study. In *Universities Transport Studies Group ANNUAL CONFERENCE*, 2019.

[146] Jinke D. Van Der Laan, Adriaan Heino, and Dick De Waard. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies*, 5(1):1–10, 1997.

[147] Petya Ventsislavova, David Crundall, Pedro Garcia-Fernandez, and Candida Castro. Assessing willingness to engage in risky driving behaviour using naturalistic driving footage: The role of age and gender. *International Journal of Environmental Research and Public Health*, 18(19), 2021.

[148] Petya Ventsislavova, Tova Rosenbloom, Joost Leunissen, Yishai Spivak, and David Crundall. An online hazard prediction test demonstrates differences in the ability to identify hazardous situations between different driving groups. *Ergonomics*, 65(8):1119–1137, 2022.

[149] Willem Vlakveld, Nicole van Nes, Jonathan de Bruin, Luuk Vissers, and Merle van der Kroft. Situation awareness increases when drivers have more time to take over the wheel in a level 3 automated car: A simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58:917–929, 10 2018.

[150] Tobias Vogelpohl, Matthias Kühn, Thomas Hummel, and Mark Vollrath. Asleep at the automated wheel—sleepiness and fatigue during highly automated driving. *Accident Analysis and Prevention*, 126:70–84, 5 2019.

[151] Bernhard Wandtner, Nadja Schömig, and Gerald Schmidt. Secondary task engagement and disengagement in the context of highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58:253–263, 10 2018.

[152] Wuhong Wang, Qian Cheng, Chenggang Li, Dietrich André, and Xiaobei Jiang. A cross-cultural analysis of driving behavior under critical situations: A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62:483–493, 2019.

[153] Joel S. Warm, Raja Parasuraman, and Gerald Matthews. Vigilance requires hard mental work and is stressful, 6 2008.

[154] Bradley W. Weaver and Patricia R. DeLucia. A systematic review and meta-analysis of takeover performance during conditionally automated driving. *Human Factors*, 64(7):1227–1260, 2022. PMID: 33307821.

[155] Klemens Weigl, Clemens Schartmüller, and Andreas Riener. Development of the questionnaire on non-driving related tasks (qndrt) in automated driving: revealing age and gender differences. *Behaviour and Information Technology*, 2022.

[156] Christopher Wilson, Diane Gyi, Andrew Morris, Robert Bateman, and Hiroyuki Tanaka. Non-driving related tasks and journey types for future autonomous vehicle owners. *Transportation Research Part F: Traffic Psychology and Behaviour*, 85:150–160, 2 2022.

[157] Philipp Wintersberger, Andreas Riener, Clemens Schartmüller, Anna-Katharina Frison, and Klemens Weigl. Let me finish before i take over: Towards attention aware device integration in highly automated vehicles. In *Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*, pages 53–65, 2018.

[158] Yanbin Wu, Ken Kihara, Kunihiro Hasegawa, Yuji Takeda, Toshihisa Sato, Motoyuki Akamatsu, and Satoshi Kitazaki. Age-related differences in effects of non-driving related tasks on takeover performance in automated driving. *Journal of Safety Research*, 72:231–238, 2 2020.

[159] Yanbin Wu, Ken Kihara, Yuji Takeda, Toshihisa Sato, Motoyuki Akamatsu, and Satoshi Kitazaki. Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving. *Accident Analysis and Prevention*, 124:202–209, 2019.

[160] Glenn Wylie and Alan Allport. Task switching and the measurement of "switch costs". *Psychological research*, 63:212–233, 2000.

[161] Glenn Wylie and Alan Allport. Task switching and the measurement of "switch costs". *Psychological Research*, 63(3):212–233, 2000.

[162] Jingyi Xiao and Konstadinos G. Goulias. Perceived usefulness and intentions to adopt autonomous vehicles. *Transportation Research Part A: Policy and Practice*, 161:170–185, 7 2022.

[163] Zhigang Xu, Kaifan Zhang, Haigen Min, Zhen Wang, Xiangmo Zhao, and Peng Liu. What drives people to accept automated vehicles? findings from a field experiment. *Transportation Research Part C: Emerging Technologies*, 95:320–334, 10 2018.

[164] Serge Yee, Lan Nguyen, Paul Green, Jessica Oberholtzer, and Baylee Miller. Visual, auditory, cognitive, and psychomotor demands of real in-vehicle tasks. Technical report, University of Michigan, Ann Arbor, Transportation Research Institute, 2007.

[165] Sol Hee Yoon, Young Woo Kim, and Yong Gu Ji. The effects of takeover request modalities on highly automated car control transitions. *Accident Analysis and Prevention*, 123:150–158, 2019.

[166] Hanna Yun and Ji Yang. Multimodal warning design for take-over request in conditionally automated driving. *European Transport Research Review*, 12, 12 2020.

[167] Razieh Zandieh and Ransford A. Acheampong. Mobility and healthy ageing in the city: Exploring opportunities and challenges of autonomous vehicles for older adults' outdoor mobility. *Cities*, 112, 5 2021.

[168] Noa Zangi, Rawan Srour-Zreik, Dana Ridel, Hadas Chasidim, and Avinoam Borowsky. Driver distraction and its effects on partially automated driving performance: A driving simulator study among young-experienced drivers. *Accident Analysis and Prevention*, 166:106565, 3 2022.

[169] Kathrin Zeeb, Axel Buchner, and Michael Schrauf. Is take-over time all that matters? the impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis and Prevention*, 92:230–239, 2016.

[170] Bo Zhang, Joost de Winter, Silvia Varotto, Riender Happee, and Marieke Martens. 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, 2019.

[171] Qiaoning Zhang, Xi Jessie Yang, and Lionel P. Robert. Drivers' age and automated vehicle explanations. *Sustainability (Switzerland)*, 13:1–19, 2 2021.

[172] Tingru Zhang, Da Tao, Xingda Qu, Xiaoyan Zhang, Rui Lin, and Wei Zhang. The roles of initial trust and perceived risk in public's acceptance of automated vehicles. *Transportation Research Part C: Emerging Technologies*, 98:207–220, 1 2019.

[173] Türker Özkan, Timo Lajunen, Joannes El. Chliaoutakis, Dianne Parker, and Heikki Summala. Cross-cultural differences in driving behaviours: A comparison of six countries. *Transportation Research Part F: Traffic Psychology and Behaviour*, 9(3):227–242, 2006.