Center for Connected and Automated Transportation

Lyles School of Civil Engineering

6-2022

Impacts of In-vehicle Alert Systems on Situational Awareness and Driving Performance in SAE Level 3 Vehicle Automation

Dustin Souders

Shubham Agrawal

Srinivas Peeta

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Report No. 30
Project Start Date: 01/01/2018
Project End Date: 12/31/2020
Report Date: June 2022

Impacts of In-vehicle Alert Systems on Situational Awareness and Driving Performance in SAE Level 3 Vehicle Automation

Dustin Souders

Assistant Professor
Clemson University

Shubham Agrawal

Post-doctoral Fellow
Clemson University

Srinivas Peeta

Frederick R. Dickerson Chair Professor
Georgia Institute of Technology

ACKNOWLEDGEMENTS AND DISCLAIMER

Funding for this research was provided by the Center for Connected and Automated Transportation under Grant No. 69A3551747105 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (OST-R), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Suggested APA Format Citation:

Souders, D., Agrawal, S., and Peeta, S. (2022). Impacts of In-vehicle Alert Systems on Situational Awareness and Driving Performance in SAE Level 3 Vehicle Automation, CCAT Report #30, Center for Connected and Automated Transportation, Purdue University, West Lafayette, IN.

Contact Information

Samuel Labi
3000 Kent Ave, West Lafayette, IN
Phone: 7654945926
Email: labi@purdue.edu

Srinivas Peeta
3000 Kent Ave, West Lafayette, IN
Phone: 4048942300
Email: peeta@gatech.edu

CCAT
University of Michigan Transportation Research Institute
2901 Baxter Road
Ann Arbor, MI 48152

uumtri-ccat@umich.edu
(734) 763-2498
www.ccat.umtri.umich.edu

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
30		

4. Title and Subtitle Impacts of In-vehicle Alert Systems on Situational Awareness and Driving Performance in SAE Level 3 Vehicle Automation	**5. Report Date** June 2022 **6. Performing Organization Code** N/A
7. Author(s) Dustin Souders, Shubham Agrawal, Srinivas Peeta	**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address Center for Connected and Automated Transportation, Purdue University, 550 Stadium Mall Drive, W. Lafayette, IN 47907; and University of Michigan Ann Arbor, 2901 Baxter Road, Ann Arbor, MI 48109	**10. Work Unit No.** **11. Contract or Grant No.** Contract No. 69A3551747105
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590	**13. Type of Report and Period Covered** Final Report, 1/1/2018 - 12/31/2020 **14. Sponsoring Agency Code** OST-R

15. Supplementary Notes
Conducted under the U.S. DOT Office of the Assistant Secretary for Research and Technology's (OST-R) University Transportation Centers (UTC) program.

16. Abstract
With greater vehicle automation, research on how the communication of capabilities and limitations of automated driving systems (ADSs) impacts safety needed. High profile crashes involving misuse of partial driving automation systems (L2) suggest that some drivers over-trust currently available L2 ADSs relative to their capabilities. This could worsen with L3 ADSs that might require driver intervention while being automated most of the time. In this simulator study, participants were provided introductory information via video that communicated the driver's role at different levels of automation as well as the capabilities and limitations of the simulated L3 ADS. This video ended with either an explicit reminder of the driver's responsibilities when using conditional driving automation (L3 Reminder condition) or highlighted benefits that might arrive with higher levels of driving automation (Future Benefits condition). Significant differences were found in ratings of familiarity, with participants in the Future Benefits condition reporting greater levels of familiarity over the course of the experiment (though still low) than their L3 Reminder condition counterparts, but few group differences in monitoring or take-over performance were found between conditions. The importance of hands-on practice for improving aspects of take-over performance was observed.

17. Key Words	18. Distribution Statement
Connected Autonomous Vehicles, Automated driving, Take over alerts, Driving simulation	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	37 pages	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

Table of Contents ii
List of Figures iv
List of Tables iv
1. INTRODUCTION 5
2. METHODS 7
2.1. Experimental Design 7
2.2. Participant Recruitment 7
2.3. Materials 7
2.3.1. Driving Simulator and Experimental Runs 7
2.3.2. Non-Driving Related Task 8
2.3.3. Eye Tracker and Monitoring Variables 8
2.3.4. Trust in Automation Questionnaire 9
2.3.5. Self-driving Car Acceptance Scale 9
2.4. Procedure 9
2.5. Analyses 10
3. RESULTS 12
3.1. Participants 12
3.2. Subjective Attitudes Towards Automation 12
3.2.1. Trust in Automation (TiA) Questionnaire 12
3.2.2. Self-driving Car Acceptance Scale (SCAS) 14
3.3. Visual Monitoring of Road 18
3.3.1. Mean Glance Duration 18
3.3.2. Attention Ratio 19
3.3.3. Glance Frequency 20
3.3.4. Time to First Glance 21
3.4. Take-over Performance 22
3.4.1. Minimum TTC 22
3.4.2. Take-over Time 23
3.4.3. Maximum Steering 24
3.4.4. Maximum Deceleration 25
4. DISCUSSION 27
4.1. Subjective Trust and Acceptance Attitudes 27

4.2. Monitoring Behavior 28
4.3. Take-over Performance 29
4.4. Limitations 29
5. CONCLUSION 31
6. ADDITIONAL ACKNOWLEDGMENTS 31
7. SYNOPSIS OF PERFORMANCE INDICATORS 32
8. OUTPUTS, OUTCOMES, AND IMPACTS 32
8.1 List of research outputs (publications, conference papers, and presentations) 32
8.2 Outcomes 32
8.3 Impacts 33
REFERENCES 34

List of Figures

Figure 1. The NDRT setup....8
Figure 2. Experimental Procedure 10
Figure 3. TiA Scores over Runs by Condition....13
Figure 4. TiA Reliability/ Competence Score and Annual Miles Driven Interaction....15
Figure 5. SCAS Scores over Time by Condition....15
Figure 6. Perceived Reliability/ Trust Time by Age Interaction....16
Figure 7. Appropriateness of Automation/ Compatibility Time by Age Interaction. 17
Figure 8. Intent to Use Automation Time by Annual Miles Driven Interaction....18
Figure 9. Mean Glance Duration over Runs by Condition. 19
Figure 10. Road Attention Ratio over Runs by Condition....20
Figure 11. Glance Frequency over Runs by Condition....21
Figure 12. Time to First Glance over Runs by Condition....22
Figure 13. Minimum TTC over Runs by Condition. 23
Figure 14. Take-over Time over Runs by Condition. 24
Figure 15. Maximum Steering over Runs by Condition....25
Figure 16. Maximum Deceleration over Runs by Condition....26

List of Tables

Table 1. Experimental Driving Simulator Run Details....8
Table 2. Variable Description....10
Table 3. Descriptive Statistics by Group 12
Table 4. ANCOVA results for TiA subscales....13
Table 5. Log(Mean Glance Duration) Fixed & Random Effects....19
Table 6. Log(Attention Ratio) Fixed & Random Effects....20
Table 7. Log(Road Glance Rate) Fixed & Random Effects 21
Table 8. Log(Time To First Glance) Fixed & Random Effects....22
Table 9. Minimum TTC Fixed & Random Effects....23
Table 10. Take-over Time Fixed & Random Effects....24
Table 11. Maximum Steering Fixed Effects....25
Table 12. Maximum Deceleration Fixed & Random Effects 26

1. INTRODUCTION

Increasing levels of vehicle automation have the potential to substantially increase traffic safety over the coming years (see Litman, 2020). Unfortunately, well-publicized crashes resulting from consumer misuse of commercially available SAE level 2 ADSs suggest that some drivers over-trust these systems, sometimes with fatal consequences (e.g., National Transportation Safety Board, 2017; Krisher, 2020). This misuse is largely due to consumer ignorance on how to properly use and rely on advanced vehicle technologies, and this is exacerbated by intentional autonowashing (i.e., misrepresenting the proper level of human supervision required by a partially or semi-autonomous system; Dixon, 2020) by media and marketing. Training has long been proposed as a means of helping calibrate consumers' trust levels to the capabilities and limitations of automated systems (Lee & See, 2004). Indeed, research shows a positivity bias regarding trust and reliance when using a new automated system (Goddard et al., 2012), particularly where the user has little information on how that system functions (Dzindolet et al., 2003).

Acceptance attitudes towards a emergent technology can play a large role in how it is adopted (e.g., Rogers, 2003), and how automated vehicles (AVs) are initially presented to prospective users has been shown to affect their acceptance attitudes towards them. Nees (2016) presented an online sample two short vignettes and assessed their acceptance attitudes toward self-driving vehicles. Participants that read the 'idealized' vignette that portrayed the automation as perfect expressed significantly higher acceptance ratings than both participants in a control condition and those that read a more 'realistic' vignette where the driver played a monitoring role and occasionally had to intervene. With varying levels of automated driving (i.e., SAE levels 3-5), unrealistic performance expectations may stem from such idealized portrayals investigated by Nees (2016) or simply a misunderstanding of the differing capabilities and limitations between these levels and misattributing higher level performance to lower level ADSs. If a driver mistakenly expects all of the benefits and performance from a SAE level 4 or 5 ADS from a level 3 ADS, long-term trust in and acceptance of the technology might suffer greatly if the automation fails and they must resume the driving task.

Recent research has also investigated the role introductory information plays in how ADSs are relied upon by drivers. A video-based driving simulator experiment found that providing the driver with printed information on the ADS alone (similar to a user's manual) did not lead to improved performance in recognizing when they needed to resume vehicle control or lead to meaningful changes in the their mental models (Boelhouwer et al., 2019). Another simulator study investigating how ADSs are introduced to users manipulated trust (i.e., promoted or lowered) in an ADS through various introductory materials (video, text description, and practice drive; Körber et al., 2018). They found that participants in a trust promoted group were more likely than their trust lowered counterparts to: (1) spend less time monitoring the road for hazards and more time engaged in a non-driving related task, (2) perform riskier take-over maneuvers in the simulator, and (3) crash in an obligatory take-over situation. Importantly, their study varied the two groups' circumstances for practicing take-over: the trust promoted group received a take-over request (TOR) in a non-critical situation free of other traffic or obstacles, whereas the trust lowered group was issued at TOR in a critical take-over situation that shared features with the obligatory take-over event that all participants later encountered in the experimental drive.

The current study sought to provide two experimental conditions an equivalent introduction via a largely similar introductory video and identical practice of a non-critical take-over situation. The main difference between these groups was how the introductory video ended, with either an explicit reminder of the necessity for the driver to maintain their attention to road hazards and their readiness to take-over, or a list of safety and convenience benefits beyond what might be expected from an L3 ADS. We sought to investigate whether this manipulation led to differences in participants' subjective trust and acceptance attitudes towards automated vehicles (AVs), visual monitoring of the road for hazards, and/or their take-over performance when resuming vehicle control.

2. METHODS

2.1. Experimental Design

The experiment employed a double-blind between-subjects design with two groups: 1.) *L3 Reminder* and 2.) *Future Benefits*. Participants in both groups viewed an informational introductory video that explained the driver's role in each of the SAE levels of automation (SAE International, 2018) as well as the capabilities, limitations, and instructions on how to operate the ADS they would be using in the driving simulator. The main difference between these groups was how this introductory video ended—with either a reminder of the necessity for the driver to maintain their attention to road hazards and their readiness to take-over (L3 Reminder) or a list of safety and convenience benefits that might be achieved in the future with SAE level 4 or 5 vehicle automation (Future Benefits) intended to inflate their estimations of the L3 ADS's capabilities.

2.2. Participant Recruitment

Participants were recruited from the community through ads in a university-wide email newsletter, paper fliers at community events, and word of mouth. Participants signed up for their experiment time through the lab's scheduling website. Criteria to participate in the study included: (1) being 18 years of age or older, (2) having a valid driver's license, (3) having no predisposition to motion sickness, (4) if over the age of 64, passing a pre-screen for memory impairment (Wechsler Logical Memory Scale; Wechsler, 1997), and (5) no physical impairment that would lead to difficulty getting into/out of the driving simulator unassisted.

2.3. Materials

2.3.1. Driving Simulator and Experimental Runs

A fixed-base driving simulator developed by AVSimulation that used three screens to provide a field of view of approximately 120° was used to implement the practice drive, ADS familiarization run (Run 0), and three experimental runs (Runs 1, 2, and 3), which were created using SCANeRStudio® 1.7 software. Each experimental run took place on a four-lane divided highway with a medium level of traffic and culminated in a questionably safe event which might require the participant to resume vehicle control (see

Table 1 for a description of each event). We modeled the road's curvature on sections of I-65 from West Lafayette to Chicago to avoid generating a completely straight path. The ADS used in the study could perform longitudinal and lateral control on the highway and had a set speed of 65 mph. It additionally had the ability to pass slower moving vehicles and provided a graded warning that consisted of an uncertainty alert (i.e., a single auditory chime that denoted when the ADS was unsure if it could navigate the road scene ahead) and a TOR (i.e., three auditory chimes in quick succession that denoted that the system required the driver to resume control as soon as possible). The driving simulator provided vehicle kinematic data (refresh rate of 20 Hz) that allowed for the following take-over performance variables to be computed within a 15-second window after the participant received the uncertainty alert: *minimum time to collision* (TTC), *take-over time* (seconds elapsed between the receipt of an alert and the resumption of vehicle control), *maximum steering* (steering wheel angle), and *maximum deceleration*.

Table 1. Experimental Driving Simulator Run Details

Run	Obligatory Take-over	Critical Event Description	Alert(s) Given	Time Budget
1	No	Lateral automation late to change lane resulting in a close pass of a slower moving vehicle	Uncertainty	6 s
2	Yes	Broken down vehicle occluded by bus	Uncertainty & TOR	8 s
3	No	Longitudinal automation late to apply brake for slower moving vehicle	Uncertainty	7 s

2.3.2. Non-Driving Related Task

The non-driving related task (NDRT) intended to simulate visual/manual distraction and consisted of a number transcription task delivered by tablet that was supported by a stand that was affixed to the base of the simulator and positioned near the center console (see Figure 1). The participant completed arithmetic problems containing two six-digit numbers by entering them into a calculator interface and pressing submit, so advanced numeracy skills were not needed.

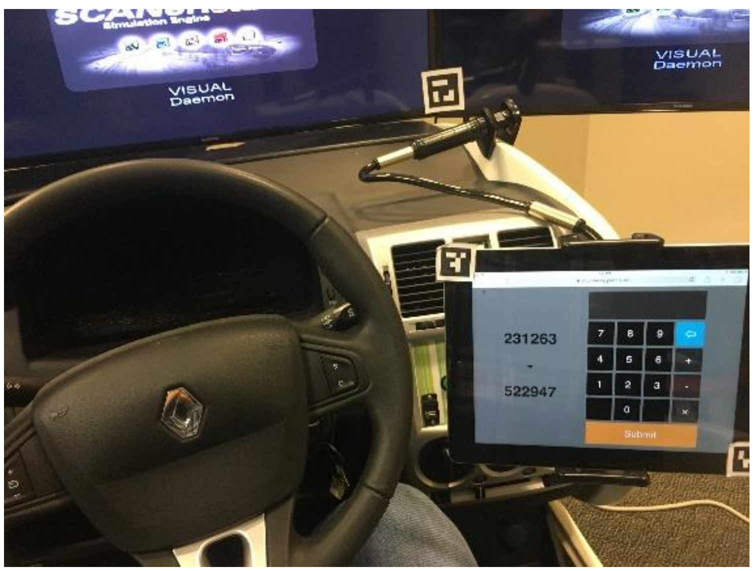

Figure 1. The NDRT setup

2.3.3. Eye Tracker and Monitoring Variables

This study used the DIKABLIS 3.0 developed by Ergoneers GmbH to track participants' eye movements. This head-mounted camera system consisted of three cameras: one forward-facing and two to track eye movements. Its design allowed for participants to wear glasses without affecting measurements. The defined areas of interest (AOIs) were *Road* (the three simulator screens) and *NDRT*. Monitoring measures gathered included *AOI Attention Ratio* (percentage of time that glances fell within the *Road* AOI), *Glance Frequency* (number of glances per unit of time when the glance was in the *Road* AOI), *Mean Glance Duration* (mean duration of all glances to the *Road* AOI), and *Time to First Glance* (elapsed time between the issuance of the uncertainty warning and the participant's first glance to the *Road* AOI).

2.3.4. Trust in Automation Questionnaire

We used an English version of the Trust in Automation (TiA) questionnaire developed and validated by Körber (2019). The version of the TiA used in the current study contained 17 items, omitting two questions regarding the "intentions of developers" subscale, and responses were made using a 5-point Likert rating scale (1 = strongly disagree, 5 = strongly agree). The remaining items broke down into the following five subscales, which were analyzed separately: Reliability/Competence, Familiarity, Trust in Automation, Understanding/Predictability, and Propensity to Trust. The TiA scale was completed by participants a total of four times: once after the ADS familiarization drive (i.e., Run 0) and each one of the three experimental runs.

2.3.5. Self-driving Car Acceptance Scale

We used an English version of the Self-driving Car Acceptance Scale (SCAS) developed and validated by Nees (2016). The SCAS consists of 24 items and responses were made using a 7-point Likert rating scale (1 = strongly disagree, 7 = strongly agree). The items broke down into 8 subscales, which were analyzed separately: Perceived Reliability/ Trust, Cost, Appropriateness/ Compatibility, Enjoyment of Driving, Perceived Usefulness, Perceived Ease of Use, Experience, and Intention to Use. The SCAS was completed by participants twice: once after their introductory video and again after the completion of the experimental runs in the driving simulator.

2.4. Procedure

The study participants, before arriving at the lab, completed a pre-experiment survey that included a screening questionnaire for pre-disposition to motion sickness, demographic information questions, and the Complacency Potential Rating Scale (CPRS; Singh et al., 1993). After the participants arrived at the lab, their informed consent was solicited, and when it was obtained, they completed a manual practice simulator drive. Following this, the experimenter left the room and participants viewed the introductory video for their randomly assigned condition. It is worth emphasizing here that both the participant and the experimenter interacting with them were both blind to condition, with a second experimenter who did not interact with the participant randomly determining the participant's condition and setting up the matching version of the introductory video. Following the introductory video, participants filled out the first of two SCASs. Next, participants were equipped with the eye tracker, electroencephalogram (EEG), and electrocardiogram (ECG). This study analyzes eye tracking data while the results for EEG and ECG data analyses are reported in other manuscripts (Agrawal & Peeta, 2021a, 2021b). Then, participants completed another practice simulator drive where they were able to practice using the NDRT and ADS concurrently. In order to ensure that participants would both engage in the NDRT as well as monitor the road, they were told that their compensation would be based on the number of NDRT trials completed as well as road safety across the three experimental runs. Participants then completed three simulator drives (approximately 10, 10, and 7 minutes, respectively) that each culminated in an event that might require them to resume control of the vehicle. In between runs, participants completed the TiA Scale (Körber, 2019). After their final run, participants completed the Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993) and a final SCAS. Participants were then debriefed and received $45 for completing the 2.5-hour study, regardless of the number of NDRT trials completed or their road safety across the three experimental runs. Figure 2 displays a flow chart of the experimental procedure.

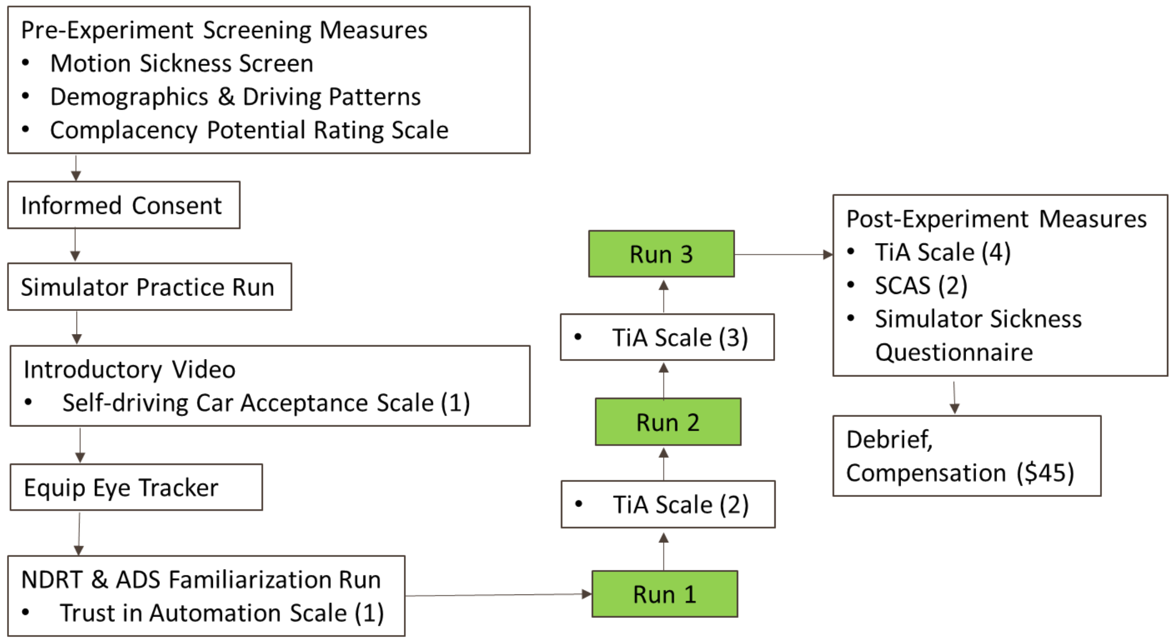

Figure 2. Experimental Procedure

2.5. Analyses

We used repeated measures ANCOVA to analyze the differences in TiA and SCAS subscales' ratings between the two groups (L3 Reminder vs. Future Benefits) over time (four TiA ratings given after each run including Run 0; two SCAS ratings given pre- and post-experiment).

Table 2. Variable Description

Variable	Definition
Condition	1 if the participant is assigned to the Future Benefits group; 0 otherwise
Age	Driver's age (in years); centered for the experiment sample
Gender	1 if female; 0 otherwise
License	1 if the participant has a driver's license for more than 5 years; 0 otherwise
Annual miles	2 if self-reported annual miles driven are more than 12,500 miles; 1 if between 7,500 and 12,499 miles; 0 otherwise
Income	1 if self-reported household annual income is more than $X; 0 otherwise
Run	Experiment run number
Time	Pre- and post-experiment time
Obligatory take-over	1 if an obligatory take-over warning was issued; 0 otherwise

For TiA analyses, the covariates included were age, gender, years with a driver's license, self-reported annual miles driven, and CPRS factor score (i.e., principle axis factoring of the trust, confidence, safety, and reliance CPRS sub-scores). Greenhouse-geisser correction was applied to within-subject factors (i.e., Run) if the assumption of sphericity was violated (i.e., p-value of Mauchly's test of sphericity is less than 0.05).

Covariates for the SCAS analyses were similar, with the only exception being the substitution of self-reported income for the CPRS factor score in order to assess if income has any effect to AV acceptance attitudes. For parsimony, only significant main effects, interactions, and covariates are reported. Covariates used for estimated marginal means in the models have the following values: Age = 28.4, Gender = 0.48, License years = 1.1, Annual miles = 1.74, CPRS factor score = -0.0005, and Income = 3.94.

We modeled the monitoring and take-over performance variables on successive runs as repeated measures using a multilevel modeling framework, including participants as a random effect in the model (using R package *lme4*; Bates et al., 2015). For each dependent variable, we included the following fixed effects: the interaction of condition and run, age (centered), whether or not take-over was obligatory (for take-over performance variables only), and participants' subjective TiA sub-score rating trust in automation (for monitoring variables only).

$$\begin{aligned}\text{Variable} \sim 1 + \text{Age} + \text{Condition} + \text{Run} + \text{Run} * \text{Condition} + \text{TiA Trust in Automation} + \\ \text{Obligatory take-over} + (1|\text{Participant})\end{aligned}$$

3. RESULTS

3.1. Participants

A total of 134 participants completed the experiment (see Table 3 for group descriptive statistics). Due to data loss for various reasons, the final sample for the TiA and SCAS analyses was 129 (65 for L3 Reminder and 64 for Future Benefits) and the final sample for the monitoring and take-over performance analyses was 132 (66 for both groups). Runs in which an unexpected rogue vehicle was present during the culminating event were removed, resulting in a total of 366 runs of a possible 396 to be analyzed.

Groups did not significantly differ on their CPRS scores ($F(1, 129) = 3.47$, $p = .07$), nor their SSQ scores ($F(1, 132) = .023$, $p = .88$). Over the course of the experiment, there were a total of three collisions (one during each run), two of which were drivers in the L3 Reminder group and one from the Future Benefits group. A two-sided Fisher's Exact test showed that this difference was not statistically significant ($p = 1.00$).

Table 3. Descriptive Statistics by Group

	L3 Reminder	Future Benefits
No. of participants	67	67
No. of females	34	36
Mean (Std. deviation) of age	28.7 (14.5)	29.0 (12.8)
No. of Take-overs across 3 runs	163	158

3.2. Subjective Attitudes Towards Automation

3.2.1. Trust in Automation (TiA) Questionnaire

Table 4 presents the analysis results of repeated measures ANCOVA for TiA subscales. Results indicate that the assumption of sphericity had been violated for all five subscales and, therefore, Greenhouse-Geisser correction is used wherever applicable. Figure 3 shows the mean scores and standard deviations of TiA subscales over runs with separate lines for both conditions. In the figure, error bars are standard deviations and dark grey line shows neutral rating point (3.0).

Table 4. ANCOVA results for TiA subscales

	Trust in Automation	Understanding/ Predictability	Familiarity	Propensity to Trust	Reliability/ Competence
Condition	-	-	$F(1, 122) = 5.09$, $p = 0.054$, $\eta_p^2 = 0.030$	-	-
Run	-	$F(2.80, 340.5) = 3.83$, $p < 0.02$, $\eta_p^2 = 0.031$	-	-	-
Run*Condition	-	-	$F(1.92, 234.6) = 6.03$, $p < 0.004$, $\eta_p^2 = 0.047$	-	-
Run*Annual miles	-	-	-	-	$F(2.64, 322) = 3.71$, $p < 0.02$, $\eta_p^2 = 0.021$
CPRS factor score	$F(1, 122) = 10.0$, $p < 0.003$, $\eta_p^2 = 0.076$	$F(1, 122) = 12.8$, $p < 0.002$, $\eta_p^2 = 0.095$	-	$F(1, 122) = 21.9$, $p < 0.001$, $\eta_p^2 = 0.152$	$F(1, 122) = 15.1$, $p < 0.001$, $\eta_p^2 = 0.110$
Mauchly's test of sphericity	$\chi^2(5) = 21.6$, $p = 0.001$	$\chi^2(5) = 14.7$, $p = 0.012$	$\chi^2(5) = 136.4$, $p < 0.001$	$\chi^2(5) = 35.9$, $p < 0.001$	$\chi^2(5) = 26.5$, $p < 0.001$

Note: Statistically non-significant results ($p > 0.05$) are not shown but were included in the model.

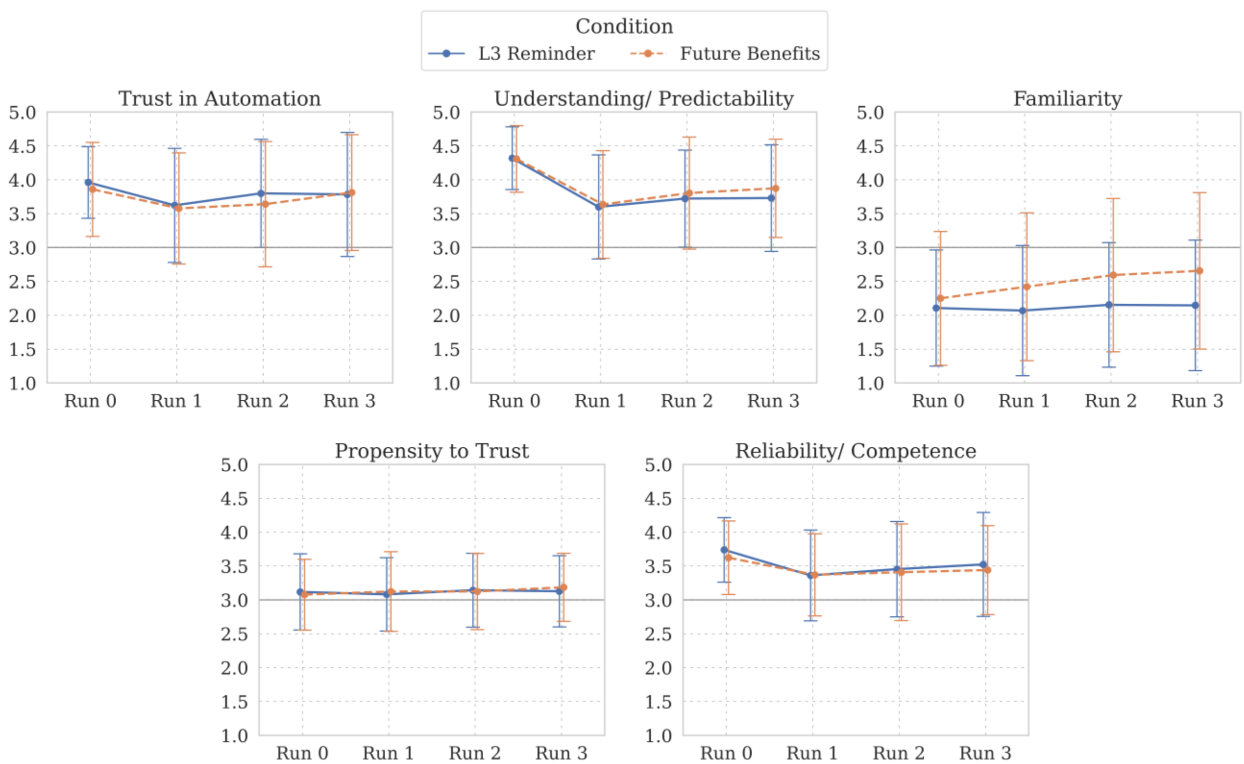

Figure 3. TiA Scores over Runs by Condition.

3.2.1.1. Trust in Automation

No within-subjects interactions or main effects were found significant. Investigating between-subjects effects, CPRS score was found to be a significant covariate ($F(1, 122) = 10.1$, $p < .003$, $\eta_p^2 = .076$), with higher CPRS scores being associated with greater trust in automation ratings.

3.2.1.2. Understanding/ Predictability

The ratings indicated that participants understood the ADS and found it predictable. The within-subjects main effect of run number was found to be significant ($F(2.80, 340.5) = 3.83$, $p < .02$, $\eta_p^2 = .031$), suggesting that ratings of understanding/ predictability differed from each other between runs. Examination of the pairwise comparisons shows that ratings after all three experimental runs were significantly lower than after Run 0, and Runs 2 and 3 were significantly higher than Run 1 (Stats???), showing that understanding/ predictability ratings steeply dropped after Run 1 and began to recover and plateau in Runs 2 & 3. Turning to the between-subject effects, CPRS factor score was found to be a significant covariate ($F(1, 122) = 12.8$, $p < .002$, $\eta_p^2 = .095$), with higher CPRS scores associated with higher ratings of understanding/ predictability.

3.2.1.3. Familiarity

The rating values suggested that participants were largely unfamiliar with systems similar to the ADS. The within-subjects interaction between run and condition was found to be significant ($F(1.92, 234.6) = 6.03$, $p < .004$, $\eta_p^2 = .047$), suggesting that ratings of familiarity differed by condition across time points. Further examination of Figure 4 shows that the future benefits group gave significantly higher ratings of familiarity over time. Investigating between-subjects effects showed that condition was marginally significant ($F(1, 122) = 5.09$, $p = .054$, $\eta_p^2 = .030$) with the future benefits condition reporting higher familiarity ratings than the L3 reminder condition, though these ratings were still in the unfamiliar range (i.e., less than the neutral point of 3) but approaching neutral for the future benefits condition.

3.2.1.4. Propensity to Trust

The ratings showed that participants were on the positive side of neutral with regards to propensity to trust the ADS. No within-subjects interactions or main effects were found to be significant. Investigating between-subjects effects found CPRS factor score to be a significant covariate ($F(1, 122) = 21.9$, $p < .0001$, $\eta_p^2 = .152$) with higher propensity to trust scores related to higher CPRS factor scores.

3.2.1.5. Reliability/ Competence

The ratings suggest that participants thought that the ADS was reliable and/or competent, though they declined after the first run and slowly recovered over the rest of the experiment. The interaction of run and annual miles driven was found to be significant ($F(2.64, 322) = 3.71$, $p < .02$, $\eta_p^2 = .021$), with participants that reported higher annual miles driven rating the reliability/ competence of the ADS lower over the course of the experiment (Figure 4). Investigating between-subjects effects found CPRS factor score to be a significant covariate ($F(1, 122) = 15.1$, $p < .0001$, $\eta_p^2 = .110$) with higher reliability/ competence scores related to higher CPRS factor scores.

3.2.2. Self-driving Car Acceptance Scale (SCAS)

Figure 5 shows the mean scores and standard deviations of SCAS subscales before and after the experiment with separate lines for both conditions. In the figure, error bars are standard deviations and dark grey line shows neutral rating point (4.0).

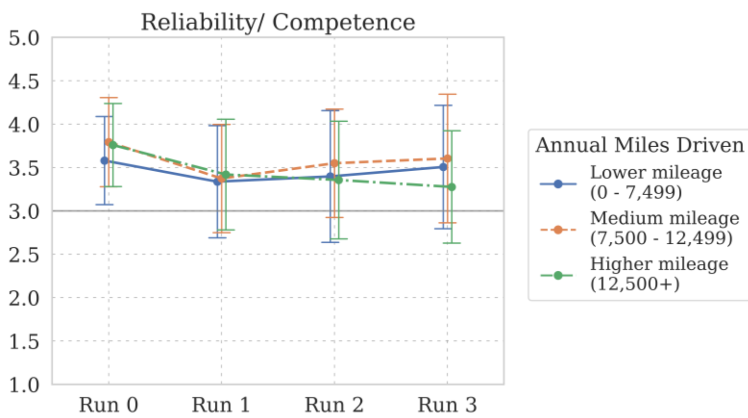

Figure 4. TiA Reliability/ Competence Score and Annual Miles Driven Interaction.

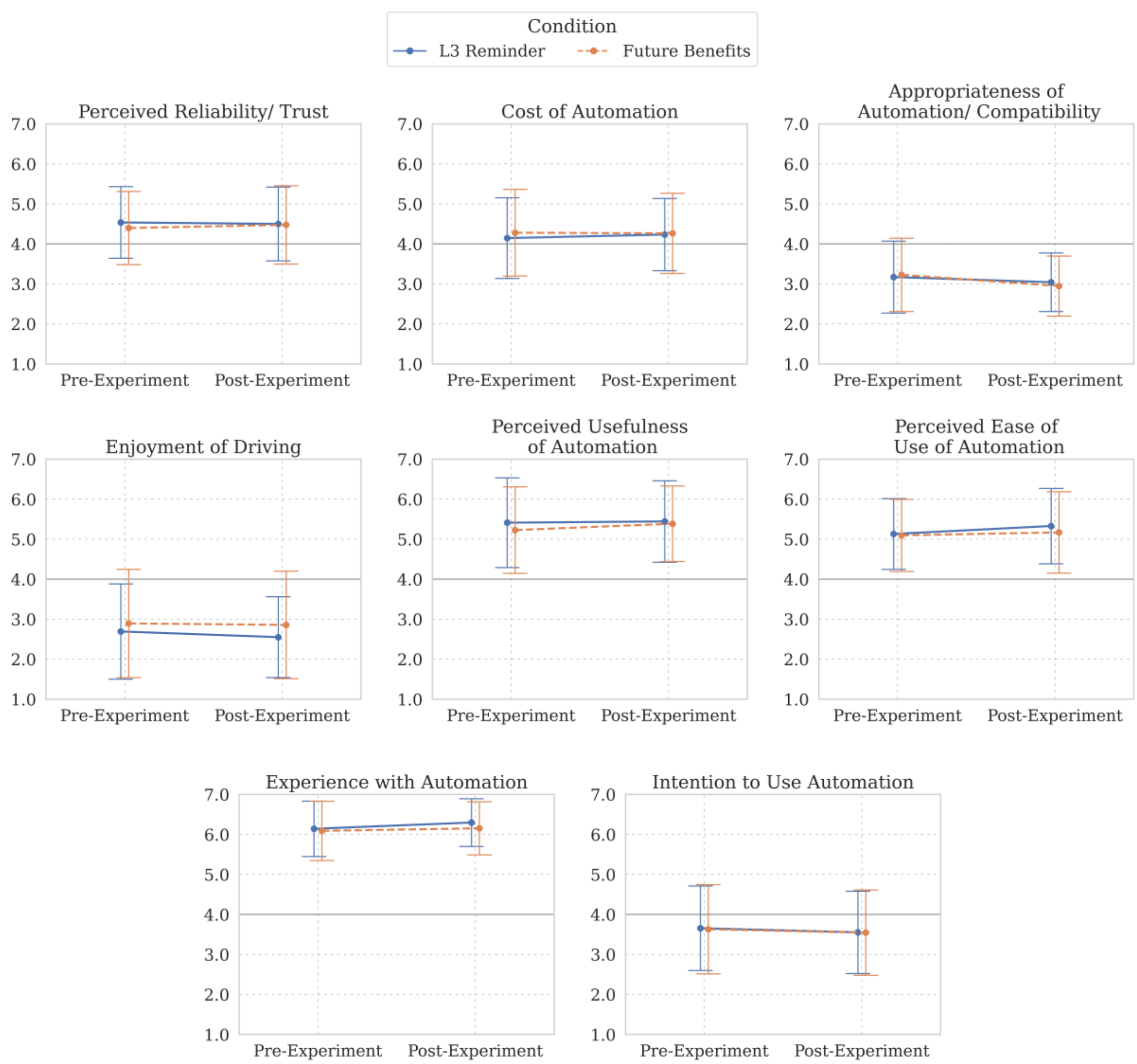

Figure 5. SCAS Scores over Time by Condition

3.2.2.1. Perceived Reliability/ Trust

The ratings hovered on the positive side of neutral with regards to pre- and post-experiment reliability/ trust and differed little over time or by condition. When assessing the within-subjects effects, the interaction of time and age was found to be marginally significant ($F(1, 122) = 3.74$, $p = .056$, $\eta_p^2 = .030$) with greater age associated with higher post-experiment perceived reliability/ trust ratings (Figure 6). No other within-subjects interactions or main effects were found to be significant. Tests of between-subjects effects showed that gender was a significant covariate ($F(1, 122) = 7.56$, $p < .008$, $\eta_p^2 = .058$), with males giving higher ratings of perceived reliability/ trust than females.

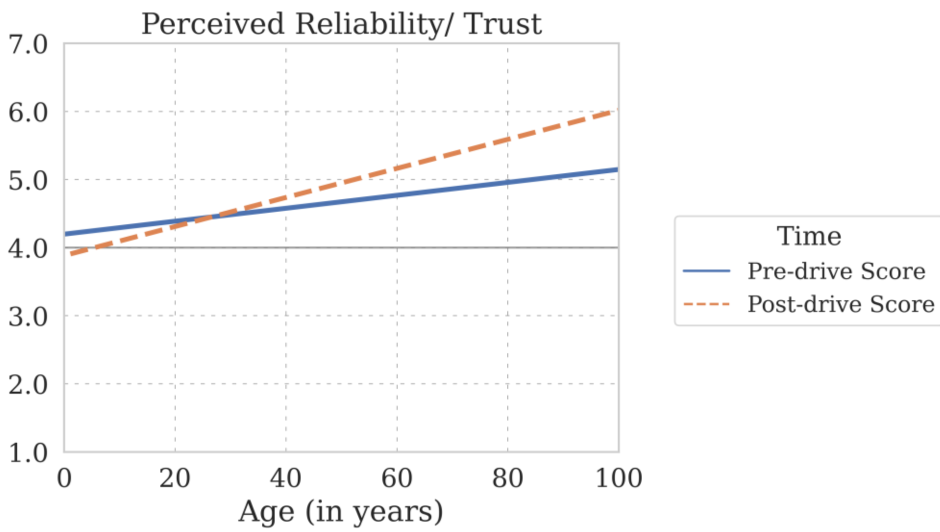

Figure 6. Perceived Reliability/ Trust Time by Age Interaction.

3.2.2.2. Cost of Automation

The ratings hovered on the positive side of neutral with regards to pre- and post-experiment cost of automation and differed little over time or by condition. No significant within-subjects interactions or main effects were found. Tests of between-subjects effects showed that gender was a significant covariate ($F(1, 122) = 5.70$, $p < .02$, $\eta_p^2 = .045$), with males expressing less cost sensitivity than females. Income was also found to be a significant covariate ($F(1, 122) = 7.34$, $p < .009$, $\eta_p^2 = .057$), with participants that reported higher annual income being less cost sensitive.

3.2.2.3. Appropriateness of Automation/ Compatibility

Though near the neutral point, ratings suggested that participants were skeptical of the appropriateness and compatibility of AVs and this differed little over time or by condition. When investigating within-subjects effects, the interaction of time and age was found to be significant ($F(1, 122) = 4.70$, $p < .04$, $\eta_p^2 = .037$) with older participants expressing higher post-experiment appropriateness of automation/ compatibility ratings (Figure 7). The tests of between-subjects effects did not yield any significant results.

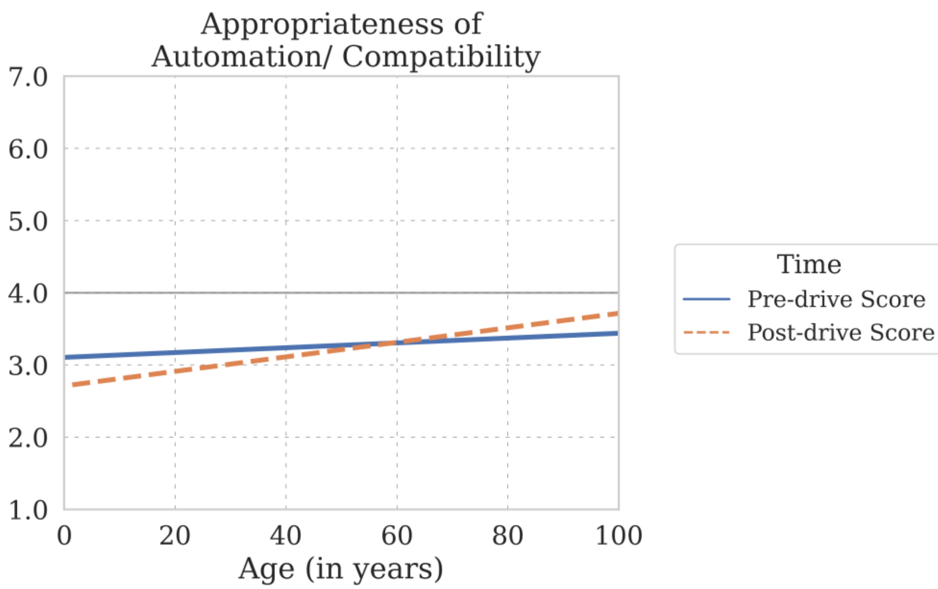

Figure 7. Appropriateness of Automation/ Compatibility Time by Age Interaction.

3.2.2.4. Enjoyment of Driving

The ratings suggest that participants do not particularly enjoy manual driving, and this differed little over time or by condition. Tests of within-subjects effects revealed no significant interactions or main effects. When investigating between-subjects effects, annual miles driven was found to be a significant covariate ($F(1, 122) = 5.19$, $p < .03$, $\eta_p^2 = .041$) with participants that reported more miles driven enjoying driving more.

3.2.2.5. Perceived Usefulness of Automation

The ratings suggest that participants found automated driving useful, and this differed little over time or by condition. Tests of within-subjects effects revealed no significant interactions or main effects. When investigating between-subjects effects, gender was found to be a marginally significant covariate ($F(1, 122) = 3.14$, $p = .08$, $\eta_p^2 = .025$) with males reporting higher perceived usefulness of automation than females.

3.2.2.6. Perceived Ease of Use of Automation

The ratings suggest that participants found automated driving easy to use, and this differed little over time or by condition. Tests of within-subjects effects revealed no significant interactions or main effects. When investigating between-subjects effects, gender was found to be a marginally significant covariate ($F(1, 122) = 3.14$, $p = .08$, $\eta_p^2 = .025$) with males reporting higher perceived usefulness of automation than females.

3.2.2.7. Experience with Automation

The ratings suggest that participants reported high levels of experience with automated technologies, and this differed little over time or by condition. Tests of within-subjects effects revealed no significant interactions or main effects. When investigating between-subjects effects, annual miles driven was found to be a marginally significant covariate ($F(1, 122) = 3.72$, $p = .056$, $\eta_p^2 = .030$) with higher mileage drivers reporting more experience with automated technologies than lower mileage drivers.

3.2.2.8. Intention to Use Automation

Hovering just below neutral, the ratings suggest that participants expressed a slight hesitance to adopt fully automated vehicles, and this differed little over time or by condition. Tests of within-subjects effects revealed a significant time*annual miles driven interaction ($F(1, 122) = 4.59$, $p < .04$, $\eta_p^2 = .036$), with higher mileage drivers' intent to use automation dropping significantly post-experiment (Figure 8). When investigating between-subjects effects, gender was found to be a significant covariate ($F(1, 122) = 5.51$, $p < .03$, $\eta_p^2 = .043$) with males reporting higher intent to use AVs than females.

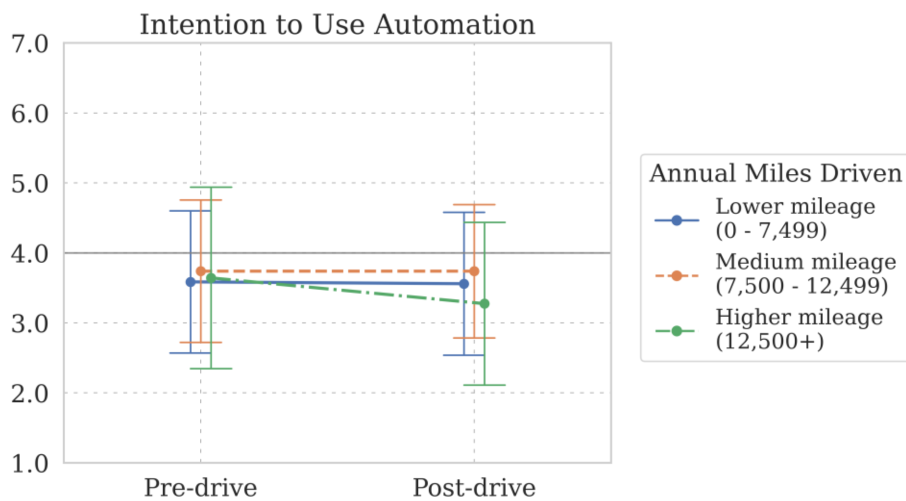

Figure 8. Intent to Use Automation Time by Annual Miles Driven Interaction.

3.3. Visual Monitoring of Road

3.3.1. Mean Glance Duration

To meet normality assumptions, mean glance duration was log-transformed. Figure 9 shows mean glance duration over runs by group and Table 5 provides information on the fixed and random effects in the mean glance duration model. The T-statistics and confidence intervals show that the effect of age was significant, with mean glance duration decreasing with greater age.

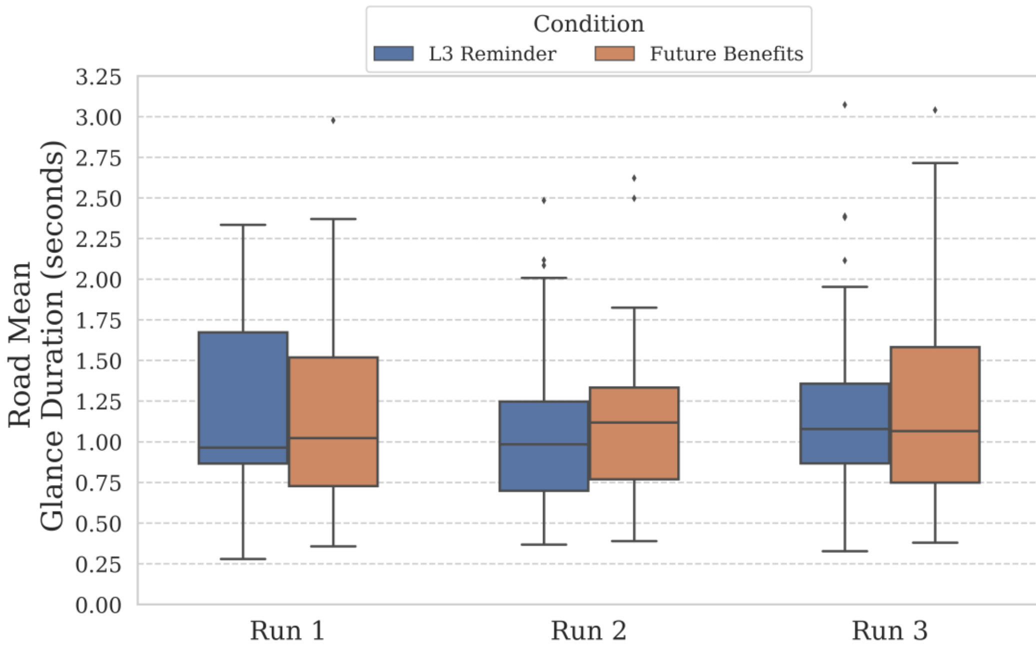

Figure 9. Mean Glance Duration over Runs by Condition.

Table 5. Log(Mean Glance Duration) Fixed & Random Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper	Random Effect (SD)
(Intercept)	0.23	0.16	1.47	-0.08	0.54	
Condition	0.01	0.09	0.17	-0.16	0.19	
Run	0.02	0.06	0.28	-0.096	0.13	0.18
Age	**-0.008**	**0.003**	**-2.59**	**-0.01**	**-0.002**	
TiA Trust in Automation	-0.06	0.04	-1.57	-0.14	0.016	
Run*Condition	-0.02	0.08	-0.24	-0.18	0.14	

Note: CI is confidence interval; Significant findings are highlighted in bold.

3.3.2. Attention Ratio

To meet normality assumptions, attention ratio was log-transformed. Figure 10 shows mean glance duration over runs by group and

Table 6 provides information on the fixed and random effects in the attention ratio model. The *T*-statistics and confidence intervals show that the effect of run was significant, with the attention ratio to the road decreasing over runs.

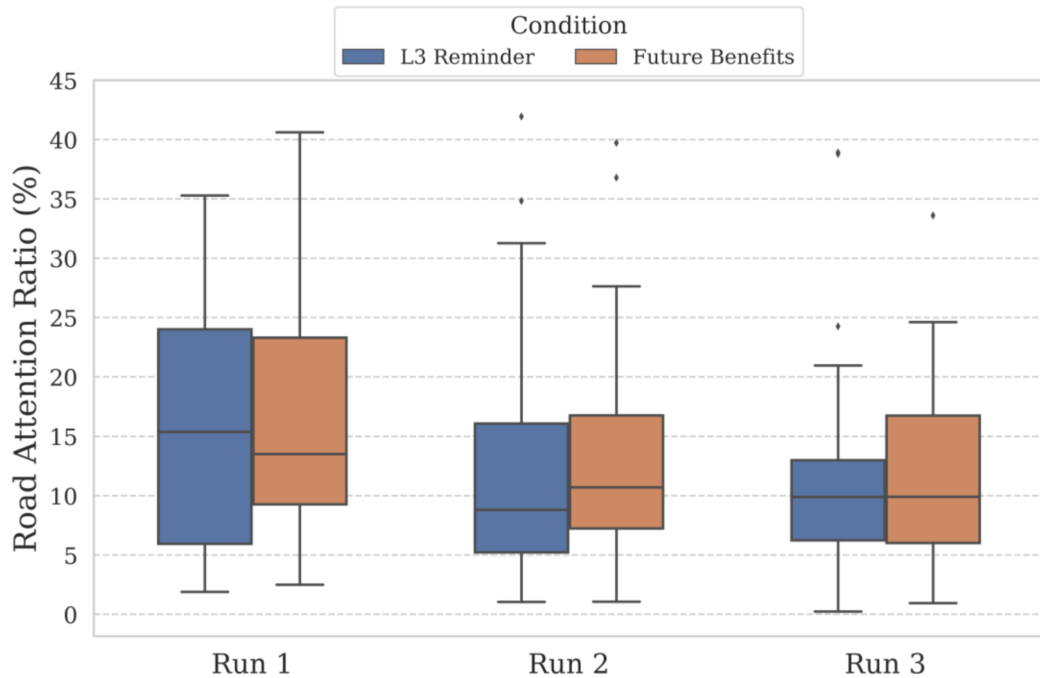

Figure 10. Road Attention Ratio over Runs by Condition.

Table 6. Log(Attention Ratio) Fixed & Random Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper	Random Effect (SD)
(Intercept)	2.58	0.1	24.8	2.38	2.79	
Condition	-0.15	0.15	-1.01	-0.44	0.14	
Run	**-0.25**	**0.1**	**-2.44**	**-0.46**	**-0.05**	**0.41**
Age	-0.001	0.01	-0.11	-0.01	0.01	
TiA Trust in Automation	-0.12	0.06	-1.95	-0.24	0.004	
Run*Condition	0.03	0.15	0.21	-0.26	0.32	

Note: CI is confidence interval; Significant findings are highlighted in bold.

3.3.3. *Glance Frequency*

To meet normality assumptions, glance frequency was log-transformed. Figure 11 shows glance frequency over runs by group and

Table 7 provides information on the fixed and random effects in the glance frequency model. The *T*-statistics and confidence intervals show that the effect of run was significant, with the glance frequency decreasing over runs.

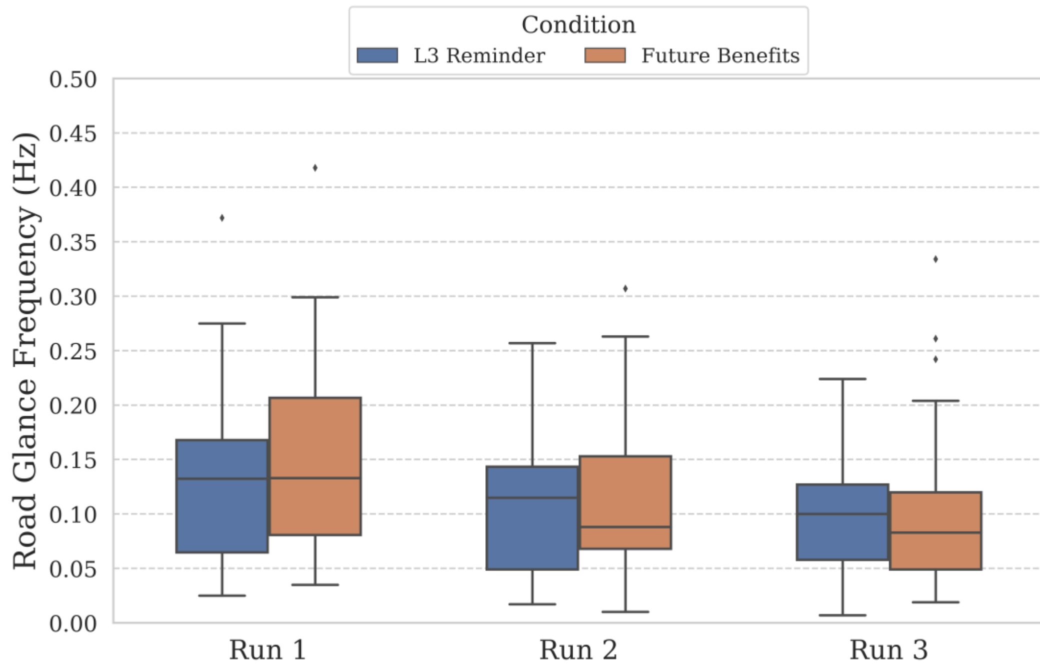

Figure 11. Glance Frequency over Runs by Condition.

Table 7. Log(Road Glance Rate) Fixed & Random Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper	Random Effect (SD)
(Intercept)	-1.8	0.22	-8.03	-2.2	-1.3	
Condition	-0.16	0.12	-1.35	-0.4	0.07	
Run	**-0.27**	**0.09**	**-3.07**	**-0.44**	**-0.1**	**0.35**
Age (centered)	0.007	0.004	1.61	-0.002	0.02	
TiA Trust in Automation	-0.08	0.06	-1.38	-0.19	0.03	
Run*Condition	0.05	0.12	0.4	-0.19	0.29	

Note: CI is confidence interval; Significant findings are highlighted in bold.

3.3.4. Time to First Glance

Time to first glance values were first filtered by the time budget for each run, with values that exceeded the time budget removed from analyses. This was largely confined to Run 3, due to a less-critical culminating event in which the ADS simply slowed to match the pace of a slower lead vehicle, and where many participants only looked up from the NDRT after they realized their speed had decreased substantially. To try to meet normality assumptions, time to first glance was log-transformed. Figure 12 shows glance frequency over runs by group and

Table 8 provides information on the fixed and random effects in the time to first glance model. The *T*-statistics and confidence intervals show that the effect of age was significant, with older participants having later first glances.

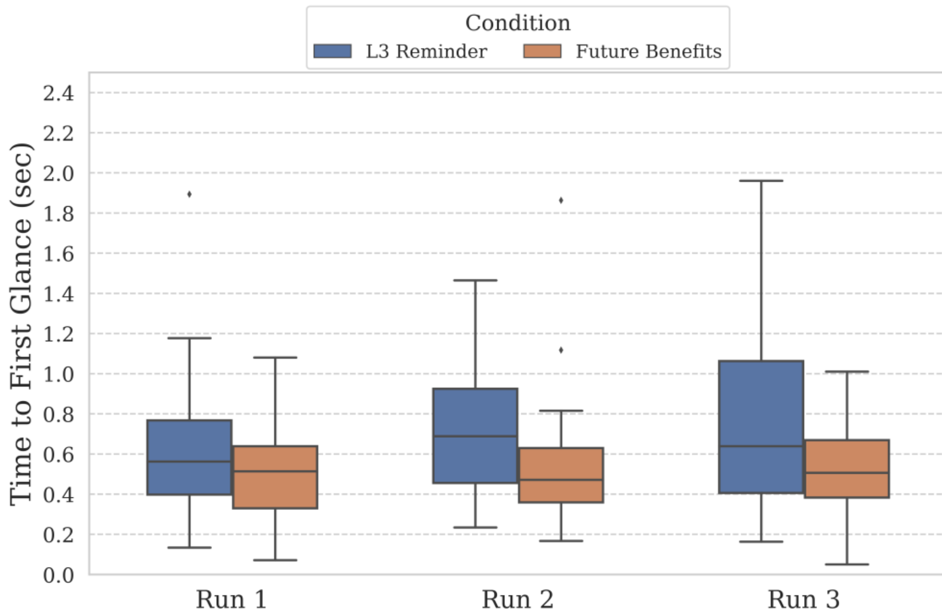

Figure 12. Time to First Glance over Runs by Condition.

Table 8. Log(Time To First Glance) Fixed & Random Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper	Random Effect (SD)
(Intercept)	-0.98	0.28	-3.54	-1.52	-0.43	
Condition	0.21	0.16	1.37	-0.09	0.52	
Run	0.02	0.11	0.19	-0.19	0.23	0.25
Age (centered)	**0.016**	**0.005**	**3.07**	**0.006**	**0.03**	
TiA Trust in Automation	0.06	0.07	0.86	-0.07	0.19	
Run*Condition	0.04	0.15	0.24	-0.25	0.32	

Note: CI is confidence interval; Significant findings are highlighted in bold.

3.4. Take-over Performance

3.4.1. Minimum TTC

Minimum TTCs were first filtered by the time budget for each run, with values that exceeded the time budget being deemed "safe take-overs" in terms of this variable and excluded from analyses. Figure 13 shows minimum TTC over runs by group and

Table 9 provides information on the fixed and random effects in the minimum TTC model. The *T*-statistics and confidence intervals show that the effects of run and obligatory take-over were significant, with minimum TTC increasing over successive runs and being lower for the obligatory take-over scenario.

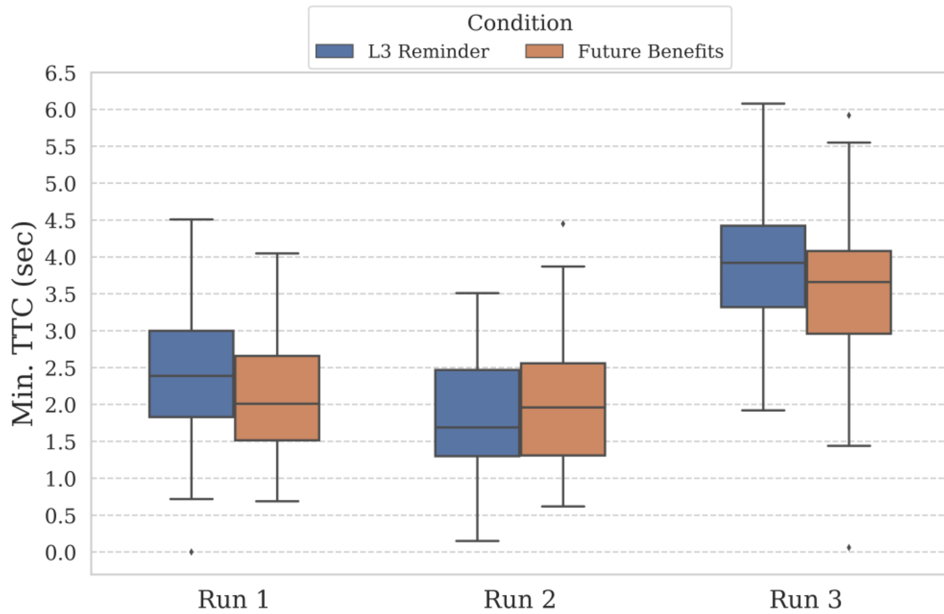

Figure 13. Minimum TTC over Runs by Condition.

Table 9. Minimum TTC Fixed & Random Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper	Random Effect (SD)
(Intercept)	2.13	0.11	19.71	1.92	2.34	
Condition	0.126	0.151	0.833	-0.17	0.42	
Run	**0.763**	**0.095**	**8.03**	**0.58**	**0.95**	**0.094**
Obligatory take-over	**-1.31**	**0.11**	**-11.85**	**-1.52**	**-1.09**	
Age (centered)	-0.003	0.004	-0.853	-0.01	0.004	
Run*Condition	-0.041	0.131	-0.31	-0.3	0.22	

Note: CI is confidence interval; Significant findings are highlighted in bold.

3.4.2. Take-over Time

Take-over times were first filtered by time budget for each run, with values that fell outside of the time budget on non-obligatory take-over runs considered unnecessary take-overs (i.e., critical event had safely passed) and excluded from analyses. Figure 14 shows take-over time over runs by condition and

Table 10 provides information on the fixed and random effects in the take-over time model. The *T*-statistics and confidence intervals show that the effects of obligatory take-over and age were significant, with longer take-over times associated with the obligatory take-over scenario and greater age.

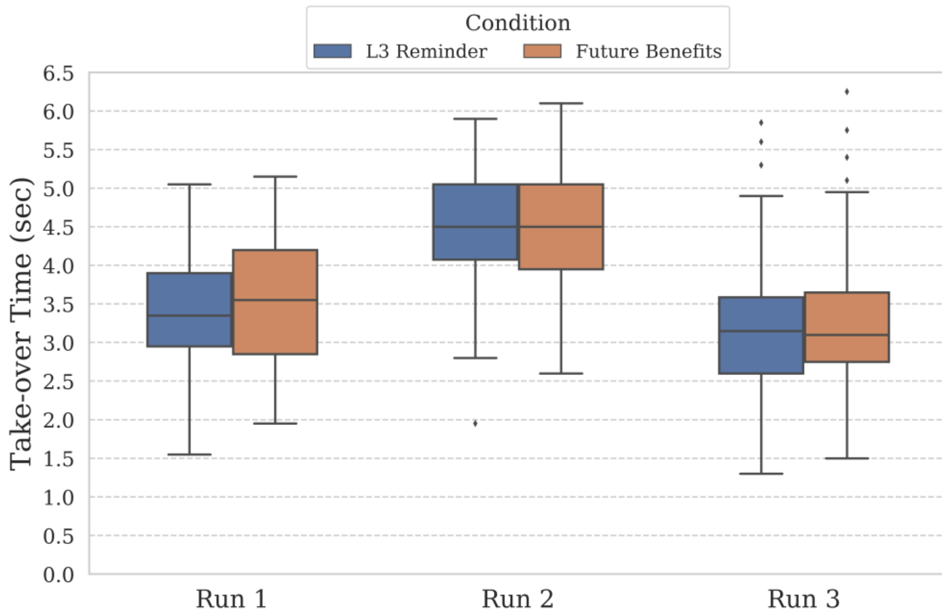

Figure 14. Take-over Time over Runs by Condition.

Table 10. Take-over Time Fixed & Random Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper	Random Effect (SD)
(Intercept)	3.51	0.09	39.73	3.34	3.68	
Condition	-0.13	0.12	-1.09	-0.38	0.11	
Run	-0.041	0.086	-0.48	-0.21	0.13	0.28
Obligatory take-over	**1.38**	**0.093**	**14.76**	**1.19**	**1.56**	
Age (centered)	**0.0086**	**0.0035**	**2.43**	**0.002**	**0.02**	
Run*Condition	0.019	0.12	0.164	-0.21	0.25	

Note: CI is confidence interval; Significant findings are highlighted in bold.

3.4.3. Maximum Steering

The multi-level model for maximum steering gave zero variance for random effects, so a linear model was run instead and reported ($F(5, 314) = 7.40$, $p < .001$, Adjusted $R^2 = .09$). Figure 15 shows maximum steering over runs by condition and

Table 11 provides information on the fixed effects in the maximum steering model. The *T*-statistics and confidence intervals show that the effects of run, obligatory take-over, and age were significant, with maximum steering values decreasing over runs, smaller during the obligatory take-over scenario and increasing with greater age.

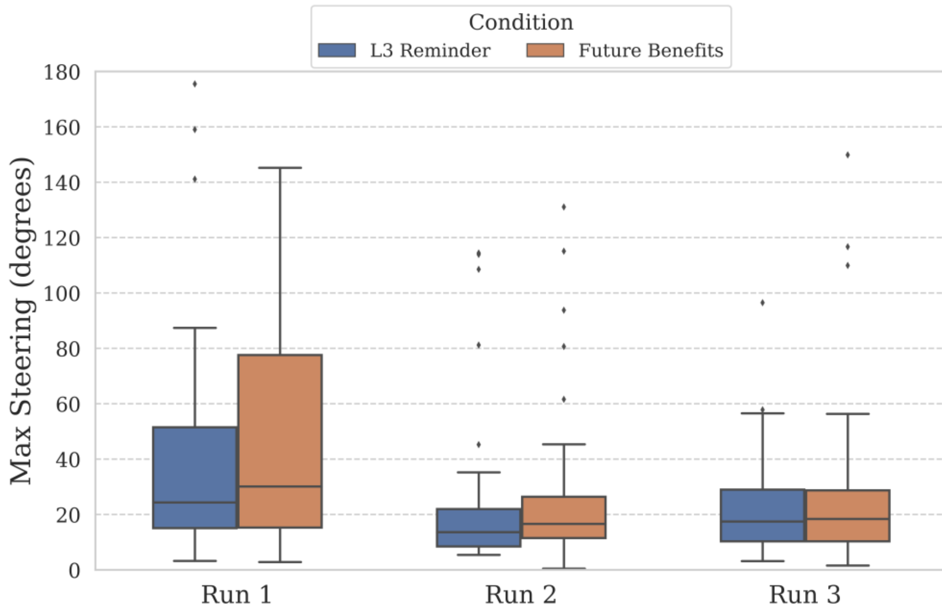

Figure 15. Maximum Steering over Runs by Condition.

Table 11. Maximum Steering Fixed Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper
(Intercept)	55.87	6.43	8.69	43.21	68.52
Condition	-0.44	8.92	-0.05	-17.98	17.1
Run	**-15.15**	**5.17**	**-2.93**	**-25.32**	**-4.97**
Obligatory take-over	**-19.57**	**6.44**	**-3.04**	**-32.24**	**-6.89**
Age (centered)	**0.73**	**0.22**	**3.34**	**0.3**	**1.16**
Run*Condition	0.83	7.26	0.11	-13.46	15.12

Note: CI is confidence interval; Significant findings are highlighted in bold.

3.4.4. Maximum Deceleration

Figure 16 shows maximum deceleration over runs by condition and

Table 12 provides information on the fixed and random effects in the maximum deceleration model. The *T*-statistics and confidence intervals show that the effects of run and obligatory take-over were significant, with maximum deceleration values decreasing over runs and being lower for the obligatory take-over scenario.

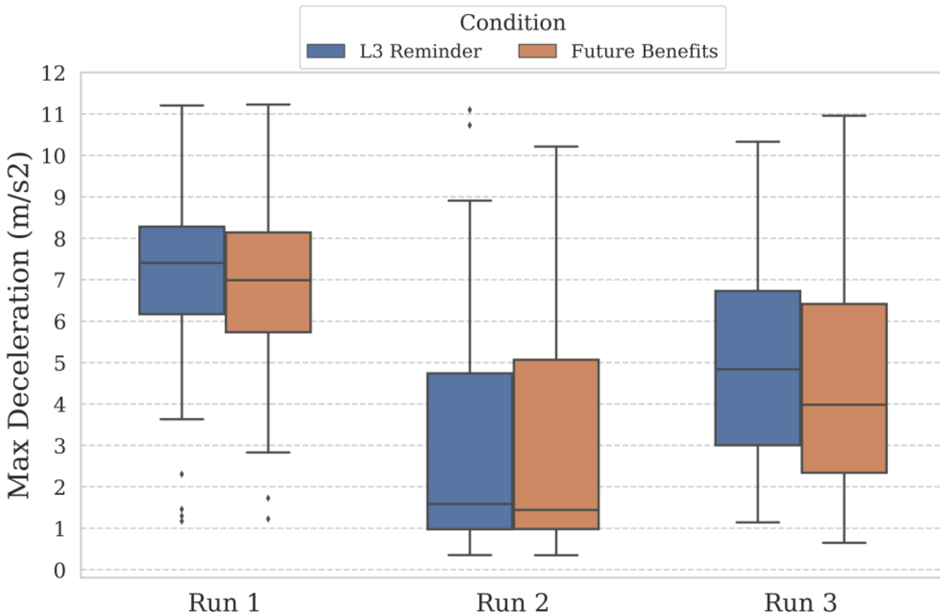

Figure 16. Maximum Deceleration over Runs by Condition.

Table 12. Maximum Deceleration Fixed & Random Effects

Term	Estimate	Std. Error	T-stat	95% CI Lower	95% CI Upper	Random Effect (SD)
(Intercept)	6.51	0.32	20.18	5.87	7.14	
Condition	0.29	0.45	0.66	-0.59	1.17	
Run	**-1.22**	**0.26**	**-4.6**	**-1.73**	**-0.7**	**0.34**
Obligatory take-over	**-2.59**	**0.32**	**-7.97**	**-3.23**	**-1.95**	
Age (centered)	0.001	0.01	0.08	-0.02	0.02	
Run*Condition	0.04	0.37	0.11	-0.69	0.77	

Note: CI is confidence interval; Significant findings are highlighted in bold.

4. DISCUSSION

Through this simulator study, we sought to investigate whether differing closing messages (i.e., a reminder of their responsibilities when using an L3 ADS or a list of benefits better associated with L4-5 ADS) to an otherwise identical informational introductory video had a differential effect on participants' trust and acceptance attitudes towards AVs, their visual monitoring of the road for hazards, and/or their take-over performance. With regard to subjective trust and acceptance attitudes, we found that participants who heard about the future benefits more associated with higher levels of automation (e.g., increased safety by removing human error, increased discretionary time while commuting) reported progressively higher levels of familiarity (i.e., that they knew or had used similar systems before) than their counterparts who were reminded of the driver's supervisory responsibilities when using an L3 ADS. While no group differences were found regarding monitoring behavior and take-over performance, participants as a whole over the course of the experiment spent less time monitoring the road for hazards, yet seemingly improved their take-over performance. These and other results and their significance are detailed in the sections below.

4.1. Subjective Trust and Acceptance Attitudes

The only significant difference found between groups in this study was a greater increase over time in the future benefits condition's TiA familiarity score. The items used to calculate this score were "I already know similar systems" and "I have already used similar systems". While the future benefits group reported numerically higher familiarity ratings at all time points and the margin between the groups widened over time, it is important to note that familiarity scores never reached the neutral point in the rating scale. The flat, somewhat stunted, familiarity ratings of the L3 reminder condition could possibly be due to a mismatch between participants' previously held ideas about automated driving where the vehicle drives itself without the need for driver supervision and/or intervention. They instead received a reminder that the L3 automation required their attention to maintain safety, whereas the future benefits condition did not, and thus avoided such a mismatch. On one hand, this finding suggests that reminding drivers that they are still involved in the driving task when using a L3 ADS might stave off any complacency that develops with increasing feelings of familiarity with an ADS. On the other hand, it might slow familiarity's beneficial effects on acceptance and adoption of ADS, as feelings of familiarity have been linked to positive attitudes towards AVs and advanced driver assistance systems (ADAS; Souders & Charness, 2016). This all might suggest that their counterparts that heard about future benefits in this study might be more complacent and more accepting, but it is worth noting the lack of significant group differences in monitoring behavior, take-over performance, and intentions to use ADS. This again could be due to the fact that familiarity attitudes for both groups were relatively low and only ever approached the neutral point of the scale.

Another novel contribution of this study were the findings around trust, acceptance, and self-reported annual miles driven. Participants who reported higher annual miles driven expressed 1.) lower ratings of reliability/competence over the course of the experiment, 2.) found the ADS easier to use, 3.) expressed less intent to use ADS after the experiment, and perhaps unsurprisingly, 4.) reported more enjoyment of driving. This constellation of findings dovetails nicely—painting a picture of seemingly competent drivers who would rather trust their own driving abilities than those of a L3 ADS due to either holding higher performance expectations, their own enjoyment of the driving task, or a mixture of both of these. Experience with similar systems also might play a factor as well, as annual miles driven was found to be a marginally significant covariate.

This study observed many significant relationships between the covariates included in the analyses and the subjective trust and acceptance attitudes assessed. Most frequent among these was the CPRS factor score, which was found to be positively related to TiA trust in automation, understanding/predictability, propensity to trust, and reliability/competence ratings. This makes sense, as the CPRS measures complacency potential with regards to a number of automated technologies (e.g., ATMs, flight automation, medical technologies) and consists of similar subscales: trust, confidence, safety, and reliance.

Regarding the demographic covariates, significant relationships were observed along the lines of gender, age, and income. Males were found to have higher levels of perceived reliability/trust, expressed marginally higher perceived usefulness, were less cost-sensitive, and displayed higher intent to use AVs than females. These findings align with most of the studies that have found a significant relationship between gender and intentions to use and/or buy AVs (e.g., Payre et al., 2014). Previous work has also found that females express greater levels of AV-related concerns than males as well (Charness et al., 2018). With regards to age, older participants in this study reported post-experiment increases in SCAS appropriateness/compatibility of automation, though these ratings were low overall. This suggests that older individuals might have increasing receptiveness to the idea of cars driving themselves more so than younger individuals after experiences like using a simulated ADS. Even though AVs have been suggested as having potential to help older adults meet their transportation needs (e.g., Reimer, 2014), most studies that have found significant effects of age on intention to use and/or buy AVs find that older age is associated with less inclination to make use of them (e.g., Haboucha et al., 2017). Finally, and perhaps unsurprisingly as other researchers have observed this (e.g., Kyriakidis et al., 2015), individuals reporting higher income were found to be less cost sensitive when it came to purchasing an AV.

4.2. Monitoring Behavior

Counter to the observed findings for take-over performance, monitoring behavior, as measured by attention ratio to the road and glance frequency, significantly *decreased* in supposed safety (i.e., less frequent and less time spent monitoring the road for hazards) over successive runs. This might be because participants acclimated to the experimental dual task of monitoring the road for hazards and completing NDRT trials (which were incentivized to ensure participants would engage in this secondary task) and were able to more effectively visually sample the road for hazards on later runs. While this finding was paired with safer take-over over successive runs, it does signal increasing complacency that might contribute to decreased safety in dissimilar scenarios than those encountered in the current study. Training programs and/or frameworks that stress the continual, periodic updating of situation awareness for threat detection, similar to those used by pilots (e.g., the aeronautical decision-making model; Federal Aviation Administration, 2009), could be used to help combat decreases in visual monitoring of the road and enhance drivers' ability to detect hazards.

This study also found that greater age was associated with shorter mean glance durations on the road as well as a slower time to first glance at the road after a warning was issued in this study. One possible explanation of these results in the current study might be that older participants spent more time engrossed in completing NDRT trials and were less likely than participants younger in age to switch between the NDRT and road monitoring tasks. Indeed, previous research shows that increased age is associated with greater difficulty maintaining and coordinating alternating between two tasks (Kray & Lindenberger, 2000). In their study of abrupt-onset hazards, Yeung & Wong (2015) found that older participants took a longer time to fixate the hazard, which aligns with our observation of older participants slower first glances to the road after warnings.

4.3. Take-over Performance

While no significant group differences between the L3 Reminder and Future Benefits conditions were found with regards to take-over performance and monitoring, this study did show that take-over safety, as measured by increasing minimum TTC, did improve over successive runs. A recent meta-analysis of take-over time studies found that mean take-over time was substantially lower when taking back control a second time (Zhang et al., 2019), which would lead to a higher minimum TTC on successive runs as observed in the current study. Extreme levels of braking and steering also improved, showing the importance of hands-on practice as they drastically reduced after participants' first experience of take-over with other traffic and obstacles. Together, these findings highlight the importance of initial exposure to take-over situations and suggest that providing practice in such situations, either in a driving simulator or on a closed course, might significantly improve the safety of drivers' maneuvers when resuming vehicle control to avoid hazards.

Greater age was associated with higher take-over time and greater steering extremes. As discussed above, previous work has shown that older drivers consistently take a longer time to fixate an abrupt-onset hazard (Yeung & Wong, 2015). With less time headway due to higher take-over time, it follows that more drastic steering responses may result.

Relative to participants in Körber et al.'s (2018) study, a greater proportion of participants in the current study from both groups resumed vehicle control when approaching a questionably safe situation. This might be due to a few reasons. One possible explanation could be that all participants received training on the driver's responsibilities when using different levels of automation and were explicitly told that the simulated ADS they would be using was L3. Another explanation could be that all participants were told that their compensation would in part be dependent on their road safety across runs, and hence participants might have been more highly motivated to resume vehicle control in *any* questionably safe situation.

4.4. Limitations

Of course, this study was not without its limitations. One limitation was the subtlety of the manipulation used to distinguish groups and the omission of a pure control condition that did not receive the SAE level training. This would have also enabled the authors to assess if teaching drivers about the SAE levels had any effect on their monitoring or take-over performance when using L3 automation. A second limitation was that the runs were always delivered in the same order, rather than randomly. This would have led practice effects to be more evenly distributed across runs. A third limitation was that the collected sample skewed towards younger adults. The age findings observed in this study make intuitive sense, but should be borne out and confirmed by future research. Lastly, the inclusion of an exit interview would have given us deeper insight into the subjective trust and acceptance attitudes observed in this study. This inclusion would have significantly added to an already lengthy experiment (2.5 hours), but as it stands, we are left to project from the ratings participants gave and postulate as to why exactly they were given.

5. CONCLUSION

In conclusion, an introductory informational video that contained either an explicit L3 reminder or listing of benefits did not make a difference in actual behavior, just feelings of familiarity. While no significant group differences between the L3 Reminder and Future Benefits conditions were found with regards to monitoring or take-over performance, this study showed that over successive runs participants resumed control with increasing time distance (i.e., higher minimum TTC) suggesting increasingly less time-critical take-over. Conversely, attention ratio to the road significantly *decreased* over the duration of the experiment. These findings highlight the potential safety gains of providing users of L3 ADSs hands-on take-over experience in a safe environment. Potential threats to safety might also be avoided with increased emphasis on periodic scanning of the road for hazards. They could also suggest new protocols and procedures that departments of motor vehicles could adopt to certify drivers for L3 automation during the transition to higher levels of vehicular automation. This might be all the more important for older drivers, who may rely more on vehicle automation.

6. ADDITIONAL ACKNOWLEDGMENTS

This work was supported by the Center for Connected and Automated Transportation University Transportation Center [grant no. 69A3551747105]. The authors would also like to thank Robert Proctor and Isis Chong for their helpful feedback when designing the experiment as well as the students from the Nextrans Center and Proctor lab for their assistance in running participants.

7. SYNOPSIS OF PERFORMANCE INDICATORS

7.1 Part I

The research from this advanced research project was disseminated to 116 people from industry, government, and academia. The research was presented at several conferences, including the 2018 INFORMS 2018Annual Meeting in Phoenix, Arizona, and the 2019 CCAT Annual Global Symposium in Ann Arbor, Michigan, 2019 International Conference on Applied Human Factors and Ergonomics, Washington D.C., and 2020 Annual Meeting of the Human Factors and Ergonomics Society (virtual). This project supported 4 students at the doctoral level. The outputs, outcomes, and impacts are described in the following sections.

7.2 Part II

Research Performance Indicators: 5 conference articles were produced from this project.

8. OUTPUTS, OUTCOMES, AND IMPACTS

8.1 List of research outputs (publications, conference papers, and presentations)

- **Souders, D.J.**, Benedyk, I., Agrawal, S., Guo, Y., & Peeta, S. (November 2018). *"Expectations of the Driver's Role when Using an Automated Driving System."* Lectern presentation at the INFORMS 2018 Annual Meeting, Phoenix, AZ.
- **Souders, D.J.** (February 2019). *"Purdue Driving Simulator Potpourri."* Lectern presentation at the Center for Connected and Automated Transportation's 2nd Annual Global Symposium, Ann Arbor, MI.
- **Souders, D.J.**, Agrawal, S., Benedyk, I., Li, Y., & Peeta, S. (July 2019). *"Tailoring Introductory Information to the Level of Automation: Effects on Attitudes and Usage of a Simulated L3 Automated Driving System."* 10th International Conference on Applied Human Factors & Ergonomics, Washington, D.C.
- **Souders, D.J.**, Agrawal, S., Li, Y., & Peeta, S. (October 2020). *"Highlighting the Driver's Responsibilities When Using Conditional Automation: Effects on Take-over Performance and Monitoring."* Virtual lectern presentation at the 64th International Annual Meeting of the Human Factors and Ergonomics Society.
- **Souders, D.J.**, Agrawal, S., Benedyk, I., Guo, Y., Li, Y., & Peeta, S. (November 2021). *"Conditional Automation and Ambiguity in the Driver's Role: Can Graded Warnings Help?"* Virtual lectern presentation at Chang'an University's College of Information Engineering's International Seminar on Cooperative and Autonomous Vehicles and Human Factors.

8.2 Outcomes

This research project sheds light on how graded alarms might help account for the unsafe introductory information that overstates conditional automation's capabilities, as well as the potential for practice in take-over situations to improve take-over performance. These findings suggest that further work should investigate the potential for graded warnings to make early automated driving systems safer despite the capacity for abuse and misuse by their users.

8.3 Impacts

Decades of Human Factors research related to the prolonged monitoring of high performing, but imperfect, automation suggest that the mixed messages conditionally automated driving systems send to drivers (i.e., the ADS is in complete control of the driving until the moment it is no longer able to be at which it alerts a likely inattentive driver) will likely lead to unsafe traffic situations. The results of the current study suggest that graded warnings that alert the driver to increased possibilities of the need for them to resume vehicle control prior to the moment where they absolutely need to resume vehicle control showed the potential to help drivers that received overestimations of their ADS' capabilities equivalently respond to dangerous take-over situations similar to drivers who received reminders of the need for them to maintain their attention to traffic conditions when using a conditionally automated ADS. Impacts include highlighting future research directions investigating the extent to which graded warnings can assist drivers in responding to vehicle automation failures, the importance of practice when resuming control from vehicle automation in potentially dangerous situations, as well as replicating a number of findings from the extant literature on drivers' ability to resume control from conditional automation that has failed.

REFERENCES

Agrawal, S., & Peeta, S. (2021a). Evaluating the impacts of situational awareness and mental stress on takeover performance under conditional automation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *83*, 210-225.

Agrawal, S., & Peeta, S. (2021b). Evaluating the impacts of driver's pre-warning cognitive state on takeover performance under conditional automation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *83*, 80-98.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boelhouwer, A., van den Beukel, A. P., van der Voort, M. C., & Martens, M. H. (2019). Should I take over? Does system knowledge help drivers in making take-over decisions while driving a partially automated car? *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 669–684. https://doi.org/10.1016/j.trf.2018.11.016

Charness, N., Yoon, J. S., Souders, D., Stothart, C., & Yehnert, C. (2018). Predictors of attitudes toward autonomous vehicles: The roles of age, gender, prior knowledge, and personality. *Frontiers in Psychology*, *9*(2589). https://doi.org/10.3389/fpsyg.2018.02589

Dixon, L. (2020). Transportation Research Interdisciplinary Perspectives Autonowashing : The Greenwashing of Vehicle Automation. *Transportation Research Interdisciplinary Perspectives*, *5*, 100113. https://doi.org/10.1016/j.trip.2020.100113

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human Computer Studies*, *58*(6), 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7

Federal Aviation Administration. (2009). *Risk Management Handbook (FAA-H-8083-2)*.

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Haboucha, C. J., Ishaq, R., & Shiftan, Y. (2017). User preferences regarding autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *78*, 37–49. https://doi.org/10.1016/j.trc.2017.01.010

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3

Körber, M. (2019). Theoretical considerations and development of a questionnaire to measure trust in automation. *Advances in Intelligent Systems and Computing*, *823*, 13–30. https://doi.org/10.1007/978-3-319-96074-6_2

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, *66*, 18–31. https://doi.org/10.1016/j.apergo.2017.07.006

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*. https://doi.org/10.1037/0882-7974.15.1.126

Krisher, T. (2020, January). *3 crashes, 3 deaths raise questions about Tesla's Autopilot*.

Kyriakidis, M., Happee, R., & De Winter, J. C. F. (2015). Public opinion on automated driving:

Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, *32*, 127–140. https://doi.org/10.1016/j.trf.2015.04.014

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50.30392

Litman, T. (2020). Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. In *Victoria Transport Policy Institute* (Vol. 42, Issue January 2014). https://doi.org/10.1613/jair.301

National Transportation Safety Board. (2017). *Driver Errors, Overreliance on Automation, Lack of Safeguards, Led to Fatal Tesla Crash*.

Nees, M. A. (2016). Acceptance of Self-driving Cars. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 1449–1453. https://doi.org/10.1177/1541931213601332

Payre, W., Cestac, J., & Delhomme, P. (2014). Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 252–263. https://doi.org/10.1016/j.trf.2014.04.009

Reimer, B. (2014). Driver Assistance Systems and the Transition to Automated Vehicles: A Path to Increase Older Adult Safety and Mobility? *Public Policy & Aging Report*, *24*(1), 27–31. https://doi.org/10.1093/ppar/prt006

Rogers, E. M. (2003). DIFFUSION of INNOVATIONS. In *Diffusion of Innovations* (5th ed.). Simon and Schuster. https://doi.org/citeulike-article-id:126680

SAE International's new standard J3016. (2018). *Levels of driving automation*. SAE International. https://doi.org/P141661

Singh, I. L., Molloy, R., & Parasuraman, R. (1993). Automation-Induced "Complacency": Development of the Complacency-Potential Rating Scale. *The International Journal of Aviation Psychology*. https://doi.org/10.1207/s15327108ijap0302_2

Souders, D., & Charness, N. (2016). Challenges of older drivers' adoption of advanced driver assistance systems and autonomous vehicles. In J. Zhou & G. Salvendy (Eds.), *International Conference on Human Aspects of IT for the Aged Population* (pp. 428–440). Springer, Cham. https://doi.org/10.1007/978-3-319-39949-2_41

Wechsler, D. (1997). Wechsler Memory Scale-Third Edition administration and scoring manual. In *The Psychological Corporation*.

Yeung, J. S., & Wong, Y. D. (2015). Effects of driver age and experience in abrupt-onset hazards. *Accident Analysis and Prevention*, *78*, 110–117. https://doi.org/10.1016/j.aap.2015.02.024

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