



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

What does an Automated Vehicle class as a hazard? Using online video-based training to improve drivers' trust and mental models for activating an Automated Vehicle

Siobhan E. Merriman^{*}, Kirsten M.A. Revell, Katherine L. Plant

Human Factors Engineering, Transportation Research Group, Faculty of Engineering and Physical Sciences, Boldrewood Innovation Campus, University of Southampton, Burgess Road, Southampton SO16 7QF, UK

ARTICLE INFO

Keywords:

Automated Vehicles
Driver Training
Mental Models
Trust in Automation
Hazard Perception

ABSTRACT

One of the arguments in favour for the introduction of Automated Vehicles (AVs) is that they will improve road safety by reducing the frequency and severity of on-road collisions. However, if drivers have a poor mental model for the capabilities and limitations of the automation, they may over-trust and activate the automation in inappropriate road conditions leading to a collision. To address this, an online video-based training programme was developed to improve drivers' mental models for when an AV can be activated, and this was compared to the current AV driver training method (i.e. owner's manual) in a matched pairs experiment. Drivers were matched on their locus of control, age and gender before reading an owner's manual (control group) or reading an owner's manual and undergoing the new online training programme (experimental group). Their trust in automation and mental models were measured before and after training. This experiment found that the online training programme in combination with an owner's manual led to a greater improvement in drivers' mental models for when the automation can be activated compared to the owner's manual in isolation. Additionally, as both training programmes exposed drivers to the limitations of the automation, both training programmes reduced drivers' trust in automation. The online training programme can be completed anywhere, at any time and on any device, which makes it highly convenient. Therefore, there could be greater acceptance amongst current licenced drivers who may not be receptive to more training.

1. Introduction

Automated Vehicles (AVs) are becoming an everyday reality with systems such as Tesla's autopilot suite (National Transportation Safety Board, 2020), BMW's collision and pedestrian warning system with brake activation (BMW, 2017) and the Mercedes Distronic Plus (Mercedes-Benz Greenwich, 2019) entering the commercial market. Considerable resources are being invested into the development of AVs due to the safety, environmental, mobility, efficiency and sustainability benefits that they are predicted to bring. For example, AVs are expected to improve road safety by reducing the frequency and severity of vehicle collisions (Stanton & Salmon, 2009; Tengilimoglu, Carsten, & Wadud, 2023), reduce vehicle pollution, fuel consumption and greenhouse gas emissions (Greenblatt & Shaheen, 2015; Bagloee, Taviana, Asadi, & Oliver, 2016) and increase mobility and reduce congestion on the road (Choi & Ji, 2015).

^{*} Corresponding author.

E-mail addresses: S.Merriman@soton.ac.uk (S.E. Merriman), K.M.Revell@soton.ac.uk (K.M.A. Revell), K.Plant@soton.ac.uk (K.L. Plant).

<https://doi.org/10.1016/j.trf.2023.08.005>

Received 12 September 2022; Received in revised form 24 May 2023; Accepted 14 August 2023

Available online 20 August 2023

1369-8478/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The Society of Automotive Engineers (SAE, 2018) defines six levels of driving automation. These range from Level 0 (No Automation) where the driver performs all driving tasks to Level 5 (Full Automation) where the automation performs all driving tasks in all road and environmental conditions. In Level 1 automation (Driver Assistance), the driver must perform all driving tasks and monitor the road environment, however the automation can provide support for steering or braking and acceleration. In Level 2 automation (Partial Automation), the automation can control the steering, braking and acceleration of the vehicle, however the driver must monitor the vehicle and the environment and take over control of the vehicle when the limitations are reached or system failures occur. In Level 3 automation (Conditional Automation), the automation controls all driving tasks and monitors the road environment. The driver is no longer required to perform the monitoring task, however they must take over control of the vehicle when asked. In Level 4 automation (High Automation), the automation performs all driving tasks and monitors the road environment. The driver is no longer required to perform takeovers as the vehicle can transition to a minimal risk condition (e.g. enter hard shoulder, safe stop). Levels 1 and 2 automation systems can be purchased on the vehicle market (e.g. Tesla's autopilot suite, see above), Level 5 AVs are a long way off, some vehicle manufactures are bypassing the development of Level 3 AVs due to their unsafe nature (e.g. Ford Motor Company, 2016; Volvo Cars, 2017) and some vehicle manufactures planned to introduce personal Level 4 AVs to the vehicle market by the mid-2020s (e.g. Ford Motor Company, 2016; Volvo Cars, 2017). Therefore there is a timely need by manufacturers to focus on AV systems which possess Level 4 capabilities, so this study focusses on one type of Level 4 AV system.

Level 4 AVs overcome some of the challenges associated with Levels 1–3 AVs. For example, the requirement for drivers to monitor the automation and the road environment in Level 2 AVs is challenging for drivers (e.g. Molloy & Parasuraman, 1996; Parasuraman, Molloy, & Singh, 1993). They are no longer actively involved in the driving task, so may get bored and fatigued and perform secondary tasks instead (Stanton, 2015). In Level 4 AVs, drivers are no longer required to perform the continuous monitoring task (see above), therefore this challenge is expected to be overcome in Level 4 AVs. However, Level 4 AVs may not overcome all challenges associated with these lower levels of automation. One such challenge is ensuring drivers have an appropriate mental model for the capabilities and limitations of the automation. Drivers tend to have a poor mental model of the automation's functions, capabilities and limitations and this can be detrimental when they use the automation (Ebnali, Kian, Ebnali-Heidari, & Mazloumi, 2019; Revell, et al., 2020; Saffarian, de Winter, & Happee, 2012). For example, if drivers believe that the automation is more capable than it actually is, they may over-trust and over-rely on the automation, activate the automation in inappropriate road conditions (e.g. absent lane markings, Lee & See, 2004; Korber, Baseler, & Bengler, 2018) and not take over control of the vehicle when needed (Barg-Walkow & Rogers, 2016; Cahour & Forzy, 2009), leading to a collision (Merriman, Plant, Revell, & Stanton, 2021b). In Levels 1 and 2 AVs, drivers can activate the automation at any time and in any road conditions and they must take over control of the vehicle when the limitations of the automation are reached (see above). As such, it is important for drivers to have an appropriate mental model for the capabilities and limitations of Levels 1 and 2 AVs, to stop these problematic behaviours from occurring. According to the SAE (2018) guidance, Level 4 systems should overcome these issues by not becoming available in unsafe road conditions and by reaching minimal risk conditions when the limitations of the automation are reached, however the technology may not meet these requirements. The technology is currently being developed, so it is unknown whether Level 4 AVs will have these capabilities in all road conditions, especially when they first come to market. The road environment and weather conditions are dynamic, highly variable and can change unpredictably, so it is unlikely that all road conditions and combinations of road conditions where the automation could be used and that are needed to meet these requirements will be tested (Merriman, Revell, & Plant, 2023b). As such, there is a risk that Level 4 AVs may not be able to detect all safe and unsafe road conditions, which means that they may not stop drivers from activating the automation or may not reach minimal risk conditions in all unsafe road conditions. Therefore, as with Levels 1–3 AVs, it is important for drivers to have an appropriate mental model for the capabilities and limitations of Level 4 AVs, as a backup for the activation and takeover tasks in case the automation does not respond appropriately. This article is focussing on this issue in relation to one Level 4 AV system, however this issue is also highly relevant for Levels 1–3 AVs (see above), so the work described in this article can also be applied to these other levels of automation. Finally, although a poor mental model can affect both activation and takeover behaviours (see above), as drivers are in full control of activating the automation, this article is focussing on drivers' mental models for when the automation can be activated. However, as with the different levels of automation, the work conducted in this article could also be applied to the takeover task.

1.1. Links between Mental Models, Trust and Behaviour

There are strong links between drivers' mental models, their trust in automation and their behaviour (Merriman, Plant, Revell, & Stanton, 2021a). A mental model is a person's knowledge and understanding of the physical world, the behaviour of a system or the automation (Stanton & Young, 2005; Saffarian, et al., 2012). Inappropriate mental models can lead to inappropriate trust levels and behaviour when operating the automation. For example, if drivers believe that the automation is more capable than it actually is, they may over-trust and over-rely on the automation (e.g. use the automation in situations which exceed the automation's capabilities: Lee & See, 2004; see above). However, if drivers believe that the automation is less capable than it actually is, they may distrust and under-rely on the automation and not use it when it is safe and appropriate to use (Lee & See, 2004; Koustanai, Cavallo, Delhomme, & Mas, 2012; Korber, et al., 2018; Boelhouwer, van den Beukel, van der Voort, & Martens, 2019). In contrast, appropriate mental models can lead to more appropriate trust levels and behaviour when operating the automation (e.g. better take over behaviours: Hergeth, Lorenz, & Krems, 2017; Korber, et al., 2018; Sportillo, Paljic, & Ojeda, 2019). Therefore, by improving (increasing the appropriateness of) drivers' mental models, drivers' will develop more appropriate trust levels and show more appropriate behaviours when operating the automation.

1.2. Current Automated Vehicle Driver Training

Training is one intervention that has been used to develop and improve trainees' mental models (e.g. Hays, Jacobs, Prince, & Salas, 1992; Krampell, Solís-Marcos, & Hjälm Dahl, 2020; Marks, Zaccaro, & Mathieu, 2000). Despite this, drivers are not given additional training for AVs other than the prerequisite owner's manual (Manser, et al., 2019) and research suggests that this may not be an effective training method. For example, drivers may not read the manual or they may skim read it, both of which reduce their understanding and knowledge (mental models) of the automation's functioning (Forster, Hergeth, Naujoks, Krems, & Keinath, 2019). Mehlenbacher, Wogalter and Laughery (2002) found that 41% (150/365) of drivers had not read their owner's manual and if they had, they only read on average half of it. Furthermore, even if drivers read an owner's manual, they may still make wrong assumptions about the automation's functioning (Cahour & Forzy, 2009), the information does not assist them with their takeover decisions and they may be unsure what the capabilities and limitations are so become surprised with the automation's actions (Boelhrouwer, et al., 2019). As such, drivers resort to trial and error (Endsley, 2017). Trial and error is an insufficient learning strategy for drivers of AVs, because mistakes and errors can have fatal consequences. Therefore, current training for AVs is insufficient and alternative solutions are needed.

1.3. Alternative Training Solutions

In a literature review of the manual and AV driver training literature, Merriman, et al. (2021a) identified 16 AV driver training studies. However, these training programmes focussed on the knowledge, skills and attitudes (KSAs) that drivers need to take over control of the vehicle (Merriman, et al., 2021a). Despite the need for drivers to have an appropriate mental model for when the automation can be activated (see section 1.1), no training study has focussed specifically on the tasks and KSAs that drivers need to activate the automation. Therefore, it is unclear what training methods may be effective in improving drivers' mental models for when the automation can be activated. However, a Training Needs Analysis (TNA) conducted by Merriman, Plant, Revell & Stanton (2023a) provided insights into potentially effective training methods. By decomposing each AV task (e.g. activation, deactivation) into its "teachable" parts, this TNA suggested that drivers need to perform three tasks to determine whether it is safe and appropriate to activate the automation; they need to (1) scan the road environment, (2) identify hazards and (3) use their knowledge (mental models) of the capabilities and limitations of the automation to determine whether the road conditions are appropriate for activation. As drivers will need to have a good situation awareness of hazardous situations, this task could be categorised as a hazard perception task (Isler, Starkey, & Sheppard, 2011). As such, current hazard perception training programmes may be an effective training method to improve drivers' mental models for when the automation can be activated, their trust in automation and behaviour when activating the automation.

1.4. Current Hazard Perception Training

In training studies and currently deployed training programmes, online methods involving hazardous images or video-clips are used to train drivers' hazard perception skills. For example, in some studies, drivers view images of hazardous scenes and highlight the hazardous areas (Pollatsek, Narayanaa, Pradhan, & Fisher, 2006) and written feedback is given for incorrect answers (Pradhan, Pollatsek, Knodler, & Fisher, 2009; Taylor, et al., 2011). In other studies, drivers watch hazardous video-clips and either click on the screen when they see a hazard (Fisher, et al., 2002) or provide commentary on the hazards that are seen (Isler, Starkey, & Williamson, 2009). Similarly, to revise for the official hazard perception test in the United Kingdom, the Driver and Vehicle Standards Agency (DVSA) recommends drivers to read the "Official DVSA Guide to Hazard Perception" and undergo practice tests (GOV.UK, n.d.). The guide combines written text, commentary clips, multiple-choice questions and multiple-choice clips to explain what hazards are, the anticipatory clues that drivers should use and where drivers should look to locate the hazards and clues (Driver & Vehicle Standards Agency, n.d.). The practice tests follow the same format as the official hazard perception test: drivers watch 14 video-clips of everyday driving scenes and click on the screen when they see a developing hazard (GOV.UK, n.d.). However written (e.g. TheoryTest.org.uk, n.d.) or graphical (e.g. circles, colours and numbers: DrivingTestSuccess, 2020) feedback is given to help drivers understand the hazard and when they should have clicked on the screen. As these hazard perception training programmes have been tried and tested, with precedence in the driver training domain, they were used to develop an online video-based training programme to improve drivers' mental models for when to activate an AV. This article will describe the development and evaluation of this new training programme.

1.5. Study Focus

The aim of this study was to investigate whether this new online video-based training programme is more effective in improving drivers' mental models for when the automation can be activated and trust in automation compared to the current training method for AVs (owner's manual). To date, no previous AV driver training study has focussed on drivers' mental models for activating the automation (Merriman, et al., 2021a). Drivers were matched on their locus of control, age and gender before reading an owner's manual or reading an owner's manual and undergoing the new online training programme. Their trust in automation and mental models were measured before and after training. Drivers' trust in automation was measured using a well-established trust questionnaire (Gold, Körber, Hohenberger, Lechner, & Bengler, 2015). Mental models cannot be measured directly, rather researchers have traditionally measured mental models by looking at the manifestation of them in trainees' behaviour and decision-making on a related task (Gentner & Stevens, 1983). As such, drivers' mental models were measured using a hazard perception-style test where higher test

scores indicated more appropriate mental models. Although mental models are strongly connected to trust and behaviour, due to the online nature of the study, drivers' behaviour could not be measured. As such, only the links between drivers' mental models and their trust in automation were investigated in this study.

1.6. Hypotheses

Both training programmes provided drivers with information about the capabilities and limitations of the automation, therefore it was expected that both training programmes would improve drivers' mental models for when the automation can be activated. Additionally, as mental models are closely connected to drivers' trust in automation (see [section 1.1](#)), it was expected that both training programmes would change drivers' trust in automation, through the improvement of their mental models. However, the extent of the improvement in drivers' mental models and the change in drivers' trust in automation may differ between the two training conditions. Therefore, the following hypotheses were made:

H₁: The type of training that drivers receive will affect the appropriateness of their mental models for when the automation can be activated.

H₂: The type of training that drivers receive will affect their trust in automation.

H₁₀: The type of training that drivers receive will not affect the appropriateness of their mental models for when the automation can be activated.

H₂₀: The type of training that drivers receive will not affect their trust in automation.

2. Method

2.1. Participants

One hundred and twelve drivers between the ages of 18 and 89 who held a valid and full UK driving licence were divided equally between the two training conditions. Twelve drivers were withdrawn from the study because they did not complete the training and second test within the timeframe required (see [section 2.3.1](#)). As such, 20 females, 29 males and 1 non-binary driver (mean age = 47.32, SD = 17.30) in the experimental group and 21 females and 29 males (mean age = 46.26, SD = 16.98) in the control group completed the study.

The two groups were closely matched on locus of control, age and gender as these variables have been shown to influence the effectiveness of driver training programmes and/or drivers' trust in automation (see [section 2.3.1](#)). Therefore, two males in the experimental condition and two females in the control condition who completed the study were excluded from the analysis because no match with another driver was found ([Stanton, Walker, Young, Kazi, & Salmon, 2007](#); [Beggiato & Krems, 2013](#)). The final sample consisted of 96 drivers, 48 in each training condition.

Key demographics for each training group are displayed in [Table 1](#). These demographic variables were measured using questionnaires, as described in [section 2.2.3](#). Chi-square tests showed that the number of males and females ($p = .578$) and total number of advanced drivers ($p = .811$) did not significantly differ between the two training groups. Independent t -tests showed that there were no significant differences between the two training groups in terms of mean age ($p = .951$), years of licence ($p = .956$), annual mileage ($p = .321$), internality score ($p = .850$) and externality score ($p = .767$). All effect sizes were small (Cohen's d ranged between 0.011 and 0.225), suggesting that statistically close participant matching was achieved ([Walker, Stanton, Kazi, Salmon, & Jenkins, 2009](#)). Ethical approval was gained by the University's Faculty Ethics Committee (Ergo: 64702). Drivers received £10 as compensation for taking part.

Table 1
Key demographics in each training condition.

Demographic		Experimental Group		Control Group	
		N		N	
Gender	Males	27		29	
	Females	20		19	
	Non-Binary	1		0	
Advanced Drivers	Males	10		10	
	Females	2		1	
	Non-Binary	0		0	
		M	SD	M	SD
Age		46.15	16.64	46.35	16.72
Years of Licence		26.90	17.04	27.10	17.31
Annual Mileage (miles)		8248.72	4627.66	9412.80	5662.90
Internality Score (min = 0, max = 75)		33.23	12.94	33.71	11.75
Externality Score (min = 0, max = 75)		24.88	10.28	25.48	9.67

2.2. Materials

2.2.1. Online Training Programme

To reflect current hazard perception training programmes (section 1.4), the online training programme consisted of two parts.

2.2.1.1. Written Information.

Firstly drivers read an overview of the SAE levels of driving automation, the capabilities of the

There are **six** levels of Automated Vehicles. These range from **No Driving Automation** (Level 0- The driver controls all driving tasks) to **Full Driving Automation** (Level 5- The automation controls all driving tasks in all road and environmental conditions). The **four** levels in between these two extremes are explained below:

Level 1- Driver Assistance:

- Driving Automation controls speed **OR** steering.
- Driver performs remaining driving tasks.
- Driver must monitor the automation and the road environment.

Level 2- Partial Driving Automation:

- Driving Automation controls speed **AND** steering.
- Driver must monitor the automation and the road environment.

Level 3- Conditional Driving Automation:

- Driving Automation controls speed **AND** steering, indicators and lights and **monitors** the environment, the vehicle and the automation.
- Driver does **NOT** need to monitor the automation or the road environment.
- Driver must take over control of the vehicle when the limitations are reached, in an emergency or when a system failure occurs.

Level 4- High Driving Automation:

- Driving Automation controls speed **AND** steering, indicators and lights and **monitors** the environment, the vehicle and the automation.
- Driver does **NOT** need to monitor the automation or the road environment.
- Driver does **NOT** need to take over control of the vehicle when the limitations are reached, in an emergency or when a system failure occurs.
- The vehicle can **automatically** drive into the hard shoulder and call roadside assistance (motorway) or slow down and come to a safe stop in the travel lane (dual carriageway).

No Automation

Full Automation

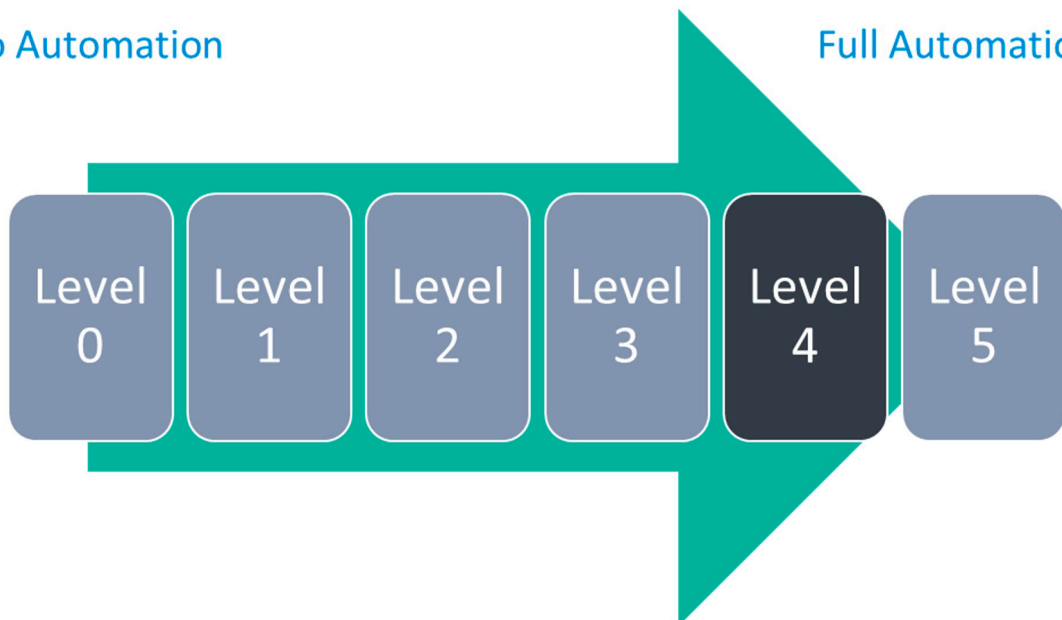


Fig. 1. Drivers read information about the SAE levels of driving automation.

automation at each level and about the focus on a Level 4 AV system (Fig. 1). Then they were given more detailed information about the capabilities and limitations of the Level 4 AV. To do this, the automation's operational design domain (road conditions where the automation is and is not designed to function: SAE, 2018) was split into five environmental categories: Road Type, Speed, Road Geometry, Weather Conditions and Roadway Conditions. For each category, drivers read about the road conditions where the automation can function reliably and the road conditions where the automation cannot function reliably (i.e. the hazards which cause that environmental category to not be satisfied). This information was based upon the operational design domain that was defined for the AV driving scenario used in Merriman, et al. (2023a). This operational design domain is summarised in Table 2. See Merriman, et al. (2023a) for more details and sources for the chosen operational design domain.

Finally, drivers read about the environmental clues (e.g. road signs, other drivers) that they can use to help them anticipate the presence of these hazards (see Fig. 2 for an example with the Road Geometry category).

2.2.1.2. Video-clip Exercises. Drivers were then shown 20 video-clips of everyday driving scenes (each one lasting between 20 and 50 seconds long). The video-clips were selected by searching YouTube and creating a list of video-clips which contained a broad range of hazards for each environmental category (e.g. roundabout, traffic lights, tractors and traffic for the speed category) and video-clips which showed appropriate road conditions (i.e. conditions where the automation can be activated). Once a lengthy list had been created, the video-clips were snipped using ShareX. Then a random number generator was used to select the video-clips for the training programme. The final list was screened and modified to ensure it represented a broad range of hazards for each environmental category.

In summary, the training video-clips represented each environmental category at least three times, with an additional three video-clips representing conditions of high/moderate automation reliability (i.e. conditions where the automation can be activated). Table 3 describes the chosen video-clips. In reality, the AV may be able to detect or predict some of these unsafe road conditions (in which case the automation would not become available). However for the purpose of this study and to ensure drivers developed an accurate mental model for when the automation can be activated, drivers were required to make all the decisions. A random number generator was used to randomise the order of the video-clips.

After viewing each video-clip, drivers were asked three questions (left-hand side of Fig. 3). Firstly, "how reliable would the automation be in these road conditions" (on a scale between 0-Highly Unreliable and 10-Highly Reliable)? Secondly, "is it safe and appropriate to activate the automation in these road conditions" (selection of "Yes" or "No")? Finally, they were asked to explain their answers to the previous two questions (open response). Once they had submitted their responses, they were given written feedback of the correct answer (e.g. right-hand side of Fig. 3). This explained what hazards they should have seen in the video-clip, how reliable the automation would have been and whether it was/was not safe and appropriate to activate the automation. This online training programme was created using the survey software Microsoft Forms.

2.2.2. Owner's Manual

To evaluate the effectiveness of this new online video-based training programme when compared to current AV driver training, a nine-page owner's manual was created (e.g. Fig. 4) as this reflects the training method that drivers currently receive for AVs (Cahour & Forzy, 2009). This manual followed the same format as current owner's manuals for Levels 1 and 2 AVs (e.g. Lexus, n.d.; Tesla, 2019; Toyota, 2019), however it was modified to reflect the greater capability of the automation (e.g. automated roadway and vehicle monitoring system) and the driver's role in Level 4 AVs (e.g. can perform secondary tasks). The owner's manual described the systems that were present, the sensors and cameras used, the capabilities and limitations of each system and the road conditions which effect the reliability of the automation and when takeover requests may occur. This owner's manual was uploaded to Microsoft Forms.

2.2.3. Questionnaires

Three questionnaires were completed during this study. The first was a demographics questionnaire which asked about the driver's age, gender, annual mileage, years of licence, advanced driving qualifications, familiarity with Advanced Driver Assistance Systems (ADAS) and current training for AVs. Advanced driving qualifications included any driving qualification that drivers had achieved after they had gained a full driving licence (e.g. Pass Plus, RoSPA or IAM RoadSmart's advanced driving qualifications). Current training for AVs included drivers who had undergone a previous driver training study for AVs, drivers who had expertise in AVs (e.g. through their employment) or drivers who had AV systems inside their current vehicle(s) and had read the owner's manual for these systems. The second was Montag and Comrey's (1987) 30-item Driving Internality and Driving Externality questionnaire. Drivers rated

Table 2
The operational design domain for the Level 4 AV that was used in this study.

Highly reliable road conditions	Moderately reliable road conditions	Highly unreliable road conditions
<ul style="list-style-type: none"> • Speed- Between 50 mph and 70 mph • Road Type- Motorways • Weather- Dry, cloudy and dull weather conditions • Roadway Conditions- Good/clear lane markings and a dry road surface • Road Geometry- Mainly straight roads 	<ul style="list-style-type: none"> • Road Type- Dual Carriageways • Weather- Light rain or moderately bright sunshine 	<ul style="list-style-type: none"> • Speed- Below 50 mph or above 70 mph • Road Type- City streets, construction zones • Weather- Heavy rain, snow, fog or bright light (from oncoming headlights or direct sunlight) • Roadway Conditions- Potholes, roadway obstacles, icy or slippery road surfaces or absent, faded or ambiguous lane markings • Road Geometry- Sharp or multiple bends

Road Geometry

The automation is **highly reliable**, so it is **safe** and **appropriate** to **activate** the automation, on **straight** (and mainly straight) roads.

The automation is **not reliable**, so it is **not safe** and **appropriate** to activate the automation, on roads which have **sharp** or **multiple bends**.

Many **clues** in the road environment can help you **anticipate** the presence of **sharp** and/or **multiple bends**. These include:

- Road **signs**.
- Route in **navigation** system
- **Behaviours** of other road users e.g. brake lights, reduction in speed, lane position
 - e.g. For a left-hand bend, position on approach is towards the offside (right) of the road.
 - For a right-hand bend, position on approach is towards the nearside (left) of the road.



Fig. 2. Example information given to drivers about the capabilities and limitations of the Level 4 AV.

Table 3

The video-clips chosen for the online video-based training programme.

Environmental Category		Description of Hazards Present
Weather Conditions	1	Heavy Rain on Dual Carriageway
	2	Bright Sunlight on Dual Carriageway
	3	Snow on Motorway
Road Type	1	End of Dual Carriageway sign
	2	Exiting Dual Carriageway
	3	Driving on a Single Carriageway
Speed	1	Traffic and restricted speed limit (40mph) on Motorway
	2	Behind a tractor on Dual Carriageway
	3	Approaching and stopping at a roundabout on Dual Carriageway
	4	Behind a cyclist on Dual Carriageway
Roadway Conditions	1	Roadworks on Motorway
	2	Poor and Absent Lane Markings on Dual Carriageway
	3	Large and deep puddles on Dual Carriageway
	4	Obstruction on information gantry on Motorway
Road Geometry	1	Sharp Bend on Motorway
	2	Multiple Bends on Dual Carriageway
	3	Sharp and Multiple Bends on Motorway
Appropriate Road Conditions	1	High Reliability on Motorway
	2	Moderate Reliability on Dual Carriageway
	3	Moderate Reliability (light rain) on Dual Carriageway

the extent to which they agreed with the 30 statements on a six-point Likert scale, ranging from 0-Disagree very much to 5-Agree very much. This questionnaire measured drivers’ locus of control and provided one score for internality and one score for externality. The sum of drivers’ scores on the internality and externality statements were used to compute their total internality and externality scores. These questionnaires were administered pre-training so that drivers could be matched on their locus of control, age and gender (see section 2.3.1). The third questionnaire was the Total Trust in Automation questionnaire (Gold, et al., 2015). This is a 35-item questionnaire which measured drivers’ trust according to six subscales: Discharge of the driver due to the automation, Safety gains, Safety hazards, Trust in automation, Perceived control of conduct and Intention to use. Drivers rated the extent to which they agreed with the statements on a five-point Likert scale (from 1-Strongly disagree to 5-Strongly agree). Drivers completed this questionnaire twice, before and after training. These questionnaires were uploaded to Microsoft Forms.

2.2.4. Tests

To measure drivers’ mental models for when the automation can be activated, drivers completed two hazard perception-style tests before and after training. These tests followed the same format as the online video-based training programme (see section 2.2.1.2); drivers were shown 20 video-clips of everyday driving scenes and were asked three questions after viewing each video-clip (e.g. Fig. 5).

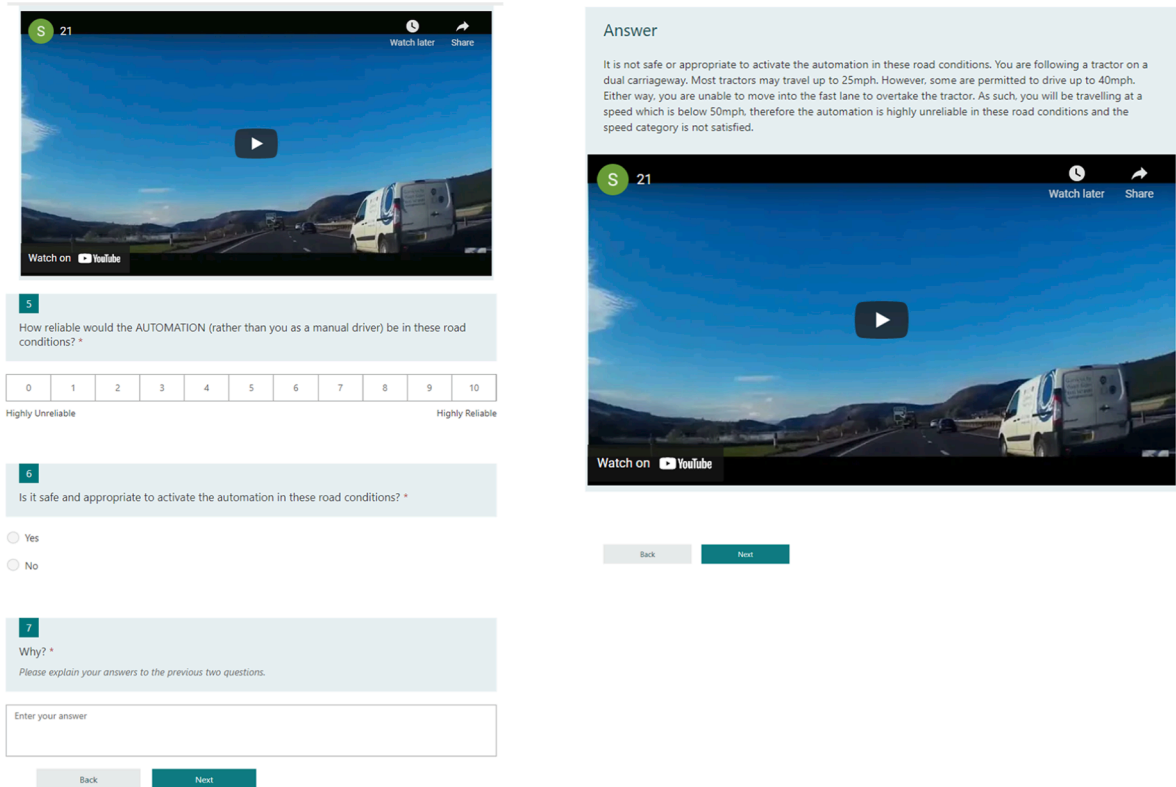


Fig. 3. An example video-clip, questions (left) and written feedback (right) in the online video-based training programme.

Automated Driving System Descriptions

Autosteer

Autosteer is an automated driving feature which automatically steers your vehicle and keeps it within the centre of your travel lane. Using various cameras and sensors, Autosteer detects lane markings and the presence of other vehicles and objects ahead to maintain the vehicle within its lane. In conjunction with the Advanced Cruise Control system, Autosteer can move the vehicle into an adjacent lane in order to overtake another vehicle.



Warning:

- Autosteer is intended for use on dual carriageways and motorways only. Do not use Autosteer on residential streets, city streets or in construction zones.
- Autosteer is designed for use on dry, straight roads which have clear lane markings, and no cross-traffic. It should not be used on roads which have very sharp turns, roads where the lane markings are absent, faded, or ambiguous or when the vehicle is being driven in a temporary or restricted lane (due to construction work).
- Autosteer is designed to work at speeds between 50 and 70mph. It should not be used at speeds below 50mph.

Fig. 4. An excerpt from the owner’s manual.

S 1 Watch later Share

Watch on YouTube

16
How reliable would the AUTOMATION (rather than you as a manual driver) be in these road conditions? *

0	1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	---	----

Highly Unreliable Highly Reliable

17
Is it safe and appropriate to activate the automation in these road conditions? *

Yes

No

18
Why? *
Please explain your answers to the previous two questions.

Enter your answer

Back Next

Fig. 5. An example video-clip and questions in the two tests.

However, unlike the online training programme, they did not receive written feedback of the correct answer.

The video-clips were selected by assigning numbers to the remaining video-clips for each environmental category (i.e. the video-clips which were not selected for the training programme) and using a random number generator to select the video-clips for each test (before and after training). The final lists were screened and modified to ensure they represented a broad range of hazards for each environmental category.

Both sets of video-clips represented each environmental category at least three times, with an additional three video-clips

Table 4
The video-clips selected for the first (before training) test.

Environmental Category		Description of Hazards Present
Weather Conditions	1	Heavy Fog on Motorway
	2	Heavy Rain on Motorway
	3	Snow on Dual Carriageway
Road Type	1	Exiting a Motorway
	2	End of Dual Carriageway sign
	3	Driving on a Single Carriageway
Speed	1	Speed Limit on Dual Carriageway is 40mph
	2	Lane Closure and slow-moving traffic on Dual Carriageway
	3	Slow-moving traffic on a Motorway
	4	Behind a slow heavy goods vehicle with no room to overtake
Roadway Conditions	1	Roadworks on Dual Carriageway
	2	Poor and Absent Lane Markings on Motorway
	3	Debris on Dual Carriageway
	4	Obstruction on information gantry on Motorway
Road Geometry	1	Sharp Bend on Motorway
	2	Sharp and Multiple Bends on Dual Carriageway
	3	Sharp Bend on Dual Carriageway
Appropriate Road Conditions	1	High Reliability on Motorway
	2	High Reliability on Motorway
	3	Moderate Reliability on Dual Carriageway

representing conditions of high/moderate automation reliability. The order of the video-clips were randomised using a random number generator. The tables below describe the video-clips that were selected for the first (before training) test (Table 4) and the second (after training) test (Table 5).

2.3. Procedure

2.3.1. Allocation of Drivers to the Training Conditions: Matching Procedure

This study took place online using Microsoft Forms. Drivers who expressed an interest to take part were sent a first Microsoft Forms link. This contained a tick box statement of consent, the demographics questionnaire, locus of control questionnaire, trust in automation questionnaire and the first (before training) test. At this first measurement point, drivers were given the SAE (2018) definition of an AV (i.e. a vehicle that is capable of performing some or all of the driving tasks that are needed to operate a vehicle on the road) and were told to base their answers on what they thought these vehicles were capable of performing.

To guarantee matching for the final few participants, the final nine drivers (9.4%) were initially screened on just the questionnaires and those who passed the screening (i.e. matched with an existing but unmatched driver) went on to complete the first test in a separate Microsoft Form.

Once the form had been submitted, the research team used the responses to match pairs of drivers together and randomly assign one driver from each pair to the experimental group (owner's manual and online training) and the other to the control group (owner's

Table 5
The video-clips selected for the second (after training) test.

Environmental Category		Description of Hazards Present
Weather Conditions	1	Fog on Motorway
	2	Heavy Rain on Motorway
	3	Bright Sunlight on Dual Carriageway
Road Type	1	Exiting a Dual Carriageway
	2	Exiting a Motorway
	3	End of Dual Carriageway sign
Speed	1	Queue and speed restriction of 40mph on information gantry on Motorway
	2	Congestion on Motorway
	3	Slow-moving traffic on Dual Carriageway
Roadway Conditions	4	Approaching and stopping at a roundabout on Dual Carriageway
	1	Poor and Absent Lane Markings on Motorway
	2	Large potholes on Motorway
	3	Lane Closure on Motorway
Road Geometry	4	Debris on Motorway
	1	Sharp Bend on Motorway
	2	Multiple Bends on Dual Carriageway
Appropriate Road Conditions	3	Sharp Bends on Dual Carriageway
	1	High Reliability on Motorway
	2	Moderate Reliability on Dual Carriageway
	3	Moderate Reliability on Dual Carriageway

manual). A matched pairs experimental design was used because this design has been successfully used in past training studies to control for variables which cannot be or are difficult to modify and are not the main focus of the study but nonetheless may have an influence on the effectiveness of the training programme and/or the system or measures used. For example [Stanton, et al. \(2007\)](#) and [Walker, et al. \(2009\)](#) matched drivers on age, driving experience, gender and annual mileage. Similarly in the AV training literature, [Cahour and Forzy \(2009\)](#) matched drivers on age and gender and [Beggiato and Krems \(2013\)](#) matched drivers on age, gender, driving style, driving experience, no ADAS experience, perceptual speed, locus of control and sensation-seeking.

In this study, drivers were matched on their locus of control, age and gender. These variables were chosen because these variables are difficult (gender, locus of control) or impossible (age) to modify but research suggests that they may influence the effectiveness of driver training programmes and/or drivers' trust in automation. For example, females ([Schoettle & Sivak, 2014](#)) and older drivers ([Lee, Ward, Raue, D'Ambrosio, & Coughlin, 2017](#)) trust AVs less and are less likely to use AVs compared to males and younger drivers. Similarly, [Stanton, et al. \(2007\)](#) found that drivers who had an internal locus of control were more proactive in trying to improve their driving skills, therefore if drivers have an internal locus of control, the training programme may be more effective. However, [Merriman, et al. \(2021\)](#) only found two papers which suggested that drivers' locus of control could be improved (i.e. internality scores increased) through training ([Huang & Ford, 2012](#); [Stanton, et al., 2007](#)). Finally, some AV driver training studies have found age and gender differences. For example, [Manser, et al. \(2019\)](#) found that females reported a greater mental effort in understanding and operating the automation when they underwent demonstration-based training. However, males reported a greater mental effort after video-based training. Similarly, [Sportillo, et al. \(2019\)](#) found that older drivers (56 years and above) reacted slowest to takeover requests when they underwent video-based training. However, their reaction times were similar to the younger drivers (below 36 years) when they underwent augmented reality training. In contrast, the younger drivers had the fastest reaction times when they underwent virtual reality training. This research demonstrates that age, gender and locus of control can influence drivers' trust in AVs, AV use and the effectiveness of driver training programmes. Therefore, by matching drivers on these three variables, this served to minimise the effect that these variables would have on the training programmes and drivers' trust scores, so that the true effect (without these confounding variables) of the training programmes on trust and mental models could be investigated ([Allen, 2017](#)).

Drivers were matched with another driver whose locus of control was in the same direction as theirs (scored higher on internality, externality or equal scores), scored within five of their internality score, within five of their externality score, whose age was within five years of their age and where possible was also the same gender (in this order). When a match was found, one driver was randomly allocated to the control condition and the other was allocated to the experimental condition. For 13 of the pairings (27%) drivers were unmatched on gender.

Within 24 hours of submitting the first form and once drivers had been allocated to one of the two training programmes, the drivers were sent a second Microsoft Forms link (or third link for the last nine drivers). This contained the training programme, the second test and the final trust in automation questionnaire. To ensure all drivers completed the study within the same time period, drivers were asked to complete this form within 48 hours of receiving the link. If they did not do this, they were withdrawn from the study ($n = 12$, see [section 2.1](#)).

2.3.2. Procedure for the Experimental Group

Upon clicking on the second Microsoft Forms link, the experimental group read the owner's manual and underwent the online video-based training programme. After a five-minute screen break, they underwent the second (after training) test and completed the trust in automation questionnaire (top level in [Fig. 6](#)).

2.3.3. Procedure for the Control Group

Upon clicking on the second Microsoft Forms link, the control group read the owner's manual. To ensure all drivers received the same amount of time for training, the control group were told that they had 45 minutes to read the owner's manual. This time was based upon the average amount of time that it took drivers to read the owner's manual and undergo the online training programme in pilot testing. Drivers could revisit the materials as many times as they wanted within this time period, however they could move onto the next task when they felt ready. After a five-minute screen break, they underwent the second (after training) test and completed the trust in automation questionnaire (bottom level in [Fig. 6](#)).

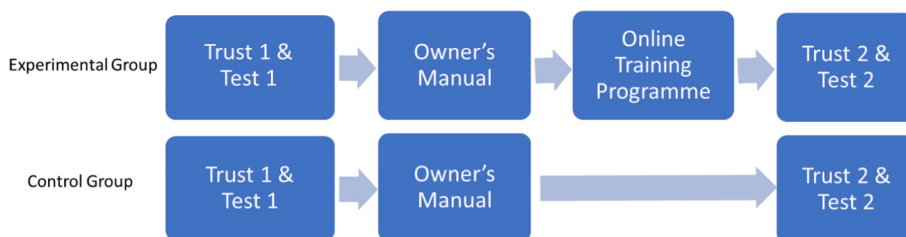


Fig. 6. Procedure for the Experimental (top) and Control Groups (bottom).

2.4. Study Design and Analysis

To analyse the effectiveness of the new online training programme when compared to current driver training for AVs (owner's manual), a matched pairs experimental design was used, whereby drivers were matched on their locus of control, age and gender before undergoing training (see [section 2.3.1](#)).

The independent variable was the type of training programme. This was a between-subjects variable which had two levels: owner's manual (control group) and owner's manual and online training (experimental group). All drivers read the owner's manual because all drivers are given an owner's manual for their vehicle (i.e. this was akin to current training provision).

There were two dependent variables. The first was the change in drivers' trust in automation from before training to after training. Drivers' trust in automation was measured using a questionnaire (see [section 2.2.3](#)). This was completed twice, before and after training. To calculate the change in each driver's trust in automation, the difference between their before and after training trust scores was calculated (after training trust score minus before training trust score).

The second dependent variable was the improvement in drivers' mental models for when the automation can be activated. Drivers' mental models were measured using a hazard perception-style test (see [section 2.2.4](#)). This was completed twice, before and after training. Although three test questions could be analysed, due to space constraints this article is focussing on the drivers' answers to the second question ("is it safe and appropriate to activate the automation in these road conditions?"). For each test, one point for the correct answer was awarded per video-clip, with a maximum score of 20 points. Therefore the higher the score, the more appropriate the drivers' mental models were for when the automation can be activated. To calculate the improvement in each driver's mental models, the difference between their before and after training test scores was calculated (after training test score minus before training test score). All variables met the assumptions for parametric testing, therefore independent measures *t*-tests were used, with type of training programme being added as the between-subjects variable. *T*-tests were performed on the difference scores, because the difference between a pre- and post-test is the typical way to measure the effects of training on the variables of interest.

Finally, as nine drivers experienced a different procedure with regards to the screening and first test (see [section 2.3.1](#)), the analysis was performed twice, once on a reduced number of matched pairs ($n = 78$, 38 pairs), so without the drivers (and their matched partner) who completed three forms and then on a full set of matched pairs ($n = 96$, 48 pairs), so with the drivers (and their matched partner) who completed three forms. As the same results were found, this analysis suggested that the change in procedure did not affect the research findings. Therefore, from this point on, this article will only report the results for the full set of matched pairs ($n = 96$, 48 pairs).

3. Results

3.1. Change in Drivers' Trust in Automation

The means and standard deviations for drivers' before and after training trust scores is displayed in [Table 6](#). This table demonstrates that generally the experimental group had higher levels of trust in automation before undergoing training. However, across both training conditions, drivers' trust in automation reduced after undergoing training. An independent measures *t*-test on the change scores revealed that there was no significant difference in the change (reduction) of drivers' trust scores from before training to after training between the two training conditions ($t(94) = 1.315$, $p = .192$, $d = 0.268$).

3.2. Improvement in Drivers' Mental Models

The means and standard deviations for drivers' before and after training mental model scores is displayed in [Table 7](#). This table demonstrates that the control group generally had more appropriate mental models for when the automation can be activated before undergoing training. Additionally, across both training conditions, drivers' mental models improved after undergoing training. An independent measures *t*-test on the improvement scores revealed that there was a significant difference in the improvement of drivers' mental models from before training to after training between the two training conditions ($t(94) = -6.919$, $p < .001$, $d = -1.412$). Drivers who read the owner's manual and underwent the online training programme showed a greater improvement in their mental models for when the automation can be activated compared to drivers who only read the owner's manual (see [Fig. 7](#)).

4. Discussion

This article describes the development of an online video-based training programme to improve drivers' mental models for when

Table 6

Means and standard deviations for drivers' trust in automation before and after undergoing training.

	Experimental Group		Control Group	
	M	SD	M	SD
Before Training Trust Score	115.92	17.14	110.63	15.55
After Training Trust Score	108.40	19.80	106.27	17.04
Change in Trust Score	- 7.52	10.61	- 4.35	12.88

Table 7

Means and standard deviations for drivers' mental models before and after undergoing training.

	Experimental Group		Control Group	
	M	SD	M	SD
Before Training Mental Model Score	9.77	3.23	11.08	3.77
After Training Mental Model Score	17.77	1.55	13.85	2.88
Improvement in Mental Model Score	8.00	3.14	2.77	4.19

A graph showing the mean improvement in drivers' mental models after undergoing one of the two types of training programmes

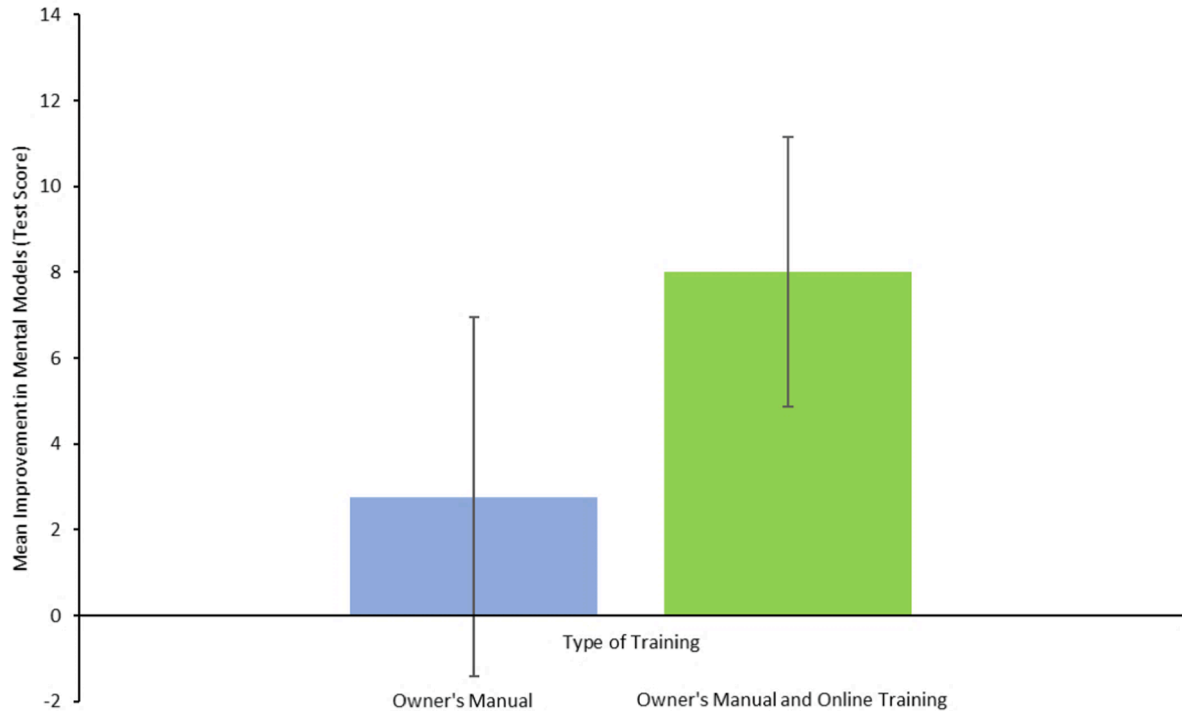


Fig. 7. The effect of type of training on the improvement of drivers' mental models (error bars represent standard deviations).

the automation can be activated and evaluates its effectiveness in a matched pairs experiment. Drivers were matched on their locus of control, age and gender before reading an owner's manual or reading an owner's manual and undergoing the new online training programme. Their trust in automation and mental models for when the automation can be activated were measured before and after training.

This study found that generally across both training conditions, drivers' trust in automation reduced after undergoing training. Although both training programmes presented information about the road conditions which were safe and unsafe for activation, a greater focus was placed on those that were unsafe. For example, in the online training programme, only three of the 20 video-clips represented appropriate (safe) road conditions. This greater focus on the unsafe road conditions could be the reason why drivers' trust in automation reduced after undergoing training. It is important for training programmes to help drivers develop appropriate levels of trust in the automation because research suggests that over- and under-trust can negatively affect drivers' behaviour. If drivers' trust is too high, they may not perform all the tasks that they are supposed to when operating the automation. For example, drivers may not monitor Level 2 AVs (Hergeth, Lorenz, Vilimek, & Krems, 2016; Rudin-Brown & Parker, 2004), may not take over control of Level 2 or 3 AVs when required (Barg-Walkow & Rogers, 2016) or may activate the automation in inappropriate road conditions (Lee & See, 2004; Korber, et al., 2018), leading to a collision (Merriman, et al., 2021b). However if drivers' trust is too low, they may not use the automation even when it is safe to use (Koustanai, et al., 2012). Therefore training programmes should not lower drivers' trust too much, because otherwise the full benefits of AVs will not be realised. Instead, they should help drivers develop appropriate levels of trust in the automation, so that they use the automation when it is safe to do so and do not use the automation when it is unsafe to do so. Therefore, future research needs to investigate whether this reduction in trust impairs or helps safe activation behaviour, and if it impairs safe activation behaviour, the online video-based training programme needs to be modified to better promote safe activation

conditions.

This study also found that the type of training that drivers received had no effect on the change (reduction) in drivers' trust in automation. This could be because both training programmes exposed drivers to the capabilities and limitations of the automation, it was just the format of the training programmes that differed (written text vs video-clips). Across both training conditions, most drivers achieved higher test scores (had more appropriate mental models) after training compared to before training (see Fig. 7), suggesting that generally drivers had become more aware of the limitations of the automation after training. To investigate this finding further, the third test question (see section 2.2.4) was analysed. Although not analysed in depth, the explanations/reasoning that drivers gave for their answers (in question three), supported the findings of question two. For example, for the video-clips which contained poor lane markings, before training 46 drivers in each condition believed that it was safe and appropriate to activate the automation, citing reasons such as "little traffic", "good visibility", "clear road", "good weather" and "no hazards". Additionally, two drivers noticed the poor lane markings (e.g. "lane markings were a little worn") but still believed that it was safe and appropriate to activate the automation. This suggests that drivers either did not see the poor lane markings and/or know that it is a hazard (limitation) for AVs. However after training, 39 drivers in the experimental condition and 24 drivers in the control condition spotted the poor lane markings and used this as reasoning for why it was not safe or appropriate to activate the automation (e.g. "road markings are faded/unclear making the automation unreliable", "poor road markings"). This suggests that both training programmes had improved drivers' mental models, and this may have reduced their trust in automation. This coincides with previous training studies which found that improved mental models of the limitations of the automation reduced drivers' trust in automation (e.g. Ebnali, et al. 2019; Hergeth, et al., 2017; Sportillo, et al., 2019).

Finally, although both training programmes improved the appropriateness of drivers' mental models, this study suggests that the combination of reading an owner's manual and undergoing the new online training programme is more effective in improving drivers' mental models than reading an owner's manual on its own. This is because drivers who read the owner's manual and underwent the online training programme showed a greater improvement in their mental models for when the automation can be activated compared to those who only read the owner's manual. Additionally, the error bars in Fig. 7 show that some drivers did worse on the hazard perception-style test (had less appropriate mental models) after reading the owner's manual. Therefore there is a greater value/benefit of undergoing the online video-based training programme, in comparison to the current training offered by vehicle manufacturers (owner's manual). This research adds to the growing literature which suggests that owner's manuals may not be the most effective training method for AVs and alternative solutions, such as simulator training (e.g. Koustanai, et al., 2012), multimedia training (e.g. Manser, et al. 2019) or online video-based training (this study) are needed.

However, this study and online training programme are not without limitations. Firstly, all drivers read the owner's manual because all drivers are given an owner's manual for their vehicle. However research shows that many drivers do not read their owner's manual and if they do, they do not read all of it (Mehlenbacher, et al., 2002). In this study, although 70% of drivers reported reading their owner's manual, the majority of drivers (37%) had read less than half and only 28% had read the whole manual. Therefore, future research should investigate the effectiveness of the online training programme in isolation (without reading an owner's manual) as this may better reflect what will occur in reality. Alternatively, mandatory reading of owner's manuals could be enforced.

Secondly, the training programmes and study were hosted online using Microsoft Forms. Although this had the advantage of being highly convenient for drivers, drivers' experience of the training programme was dependent upon the device used and the quality of their internet connection. Three drivers who underwent the online training programme reported having poor internet connection. This, along with the size of the screen, will have influenced the quality of the video-clips as the video-clips may have become blurry. Therefore, the effectiveness of the training programme on drivers' trust and mental models could have been underestimated (e.g. one driver reported being unable to read some road signs). As such, alternative methods of delivery which do not require internet connection (e.g. CD-ROM, USB stick) should be investigated so that the effectiveness of the training programme is not dependent upon these factors and all drivers undergo the same experience. Although one driver reported success when using a mobile phone, more extensive piloting should be done using different sized screens and different video-clips to investigate which video-clips are best to use. This will reduce the effect that these variables may have on the effectiveness of the training programme, so the true effect of the training programme on trust and mental models can be investigated. Additionally, this training programme was designed with the purpose of being similar to current hazard perception training programmes. However age and gender can affect how people respond to and use interfaces, how people process information and their preferences and success in using different interface designs (e.g. Hsu, 2012; Sánchez-Franco, 2006; Silva, Holden, & Nii, 2014; Watkins, Kules, Yuan, & Xie, 2014). All users of AVs will need to undergo training for AVs, so this training programme will need to be optimised for inclusivity to gain a better understanding of its effectiveness for drivers of different ages and genders.

Finally, the effectiveness of the online training programme was only evaluated in the short term (immediately after training). To evaluate the full effectiveness of the training programme, a longitudinal evaluation is needed to determine whether the improved mental models are maintained in the long term (e.g. after two weeks, a month, a year). Additionally, the transfer between drivers' improved mental models and their actual activation behaviour warrants further research (transfer of training: Baldwin & Ford, 1988; Salas & Cannon-Bowers, 2001; Krampell, et al., 2020). As mentioned in section 1.1, there are strong links between mental models, trust and behaviour. However, due to the online nature of the study, drivers' behaviour could not be measured, so only the links between drivers' mental models and trust were investigated. As such, future simulator and on-road experiments are needed to investigate how the improvement in drivers' mental models and reduction in drivers' trust impact their activation behaviour. Reliance on the automation is also measured through drivers' behaviour, so by investigating the links with drivers' behaviour, this will allow researchers to investigate over- and under-reliance which is another important issue that needs to be considered and addressed in driver training (see section 1.1). Additionally, this study involved drivers watching video-clips, so the consequences of having an inappropriate mental

model for when the automation can be activated were much lower than in real life as drivers only received no marks for an incorrect answer. In reality, if drivers have an inappropriate mental model, they may activate the automation in inappropriate road conditions and cause a collision, resulting in fatal consequences (e.g. Merriman, et al., 2021b). As the perception of risk was much lower in this study, drivers may have been riskier and said that it was safe and appropriate to activate the automation in more situations than they would actually activate the automation in real life. Therefore, to evaluate the full effectiveness of the online training programme, longitudinal studies, simulator and on-road experiments are needed to see whether these improved mental models are maintained in the long term and translate to actual on-road activation behaviour (e.g. drivers activate the automation when it is safe to do so and do not activate the automation when it is unsafe to do so).

This study is one of the first studies to develop and evaluate the effectiveness of an online training programme which targets drivers' mental models for when the automation can be activated. As the training programme is online, it is more convenient than other training methods which have been used in past training studies (e.g. simulator, virtual reality: Sportillo, et al., 2019). Drivers can complete this training programme anywhere, at any time and on any device (e.g. mobile phone, laptop, tablet). They do not have to travel to training providers which may be required for some of the other training methods (e.g. simulator, on-road). The online training programme followed the same format that is used in current hazard perception training programmes. This makes it more user-friendly than other training methods which may be unfamiliar to drivers (e.g. simulators). This familiarity and convenience may make this online training programme more acceptable to current licenced drivers, who have already undergone extensive training to learn how to drive a manual vehicle, and are not receptive to more training for AVs. Finally, the focus of this training programme was on one type of Level 4 AV system and on drivers' mental models for when the automation can be activated (i.e. from manual to AV control). However, it is also important for drivers to have an appropriate mental model for Levels 1–3 AVs, therefore this training programme may also be useful in improving drivers' mental models for when to activate Levels 1–3 AVs. Additionally, many of the scenarios that were used in the training programme also generalise to situations when the automation should be deactivated (i.e. from AV to manual vehicle control). For example, the automation is highly unreliable in heavy rain. This means that drivers should not activate the automation in heavy rain but also, if the automation has been activated and heavy rain starts to fall, the automation should be deactivated. As this training programme was effective in improving drivers' mental models for when the automation can be activated, this training programme may also be effective in improving drivers' mental models for when the automation should be deactivated. Therefore, future research should investigate the generalisability of this training programme and see whether it can also improve drivers' mental models for when the automation should be deactivated.

5. Conclusion

AVs are becoming an everyday reality. Unless drivers have an appropriate mental model for when the automation can be activated, they may not activate the automation safely or appropriately on the road. Therefore, training in the appropriate activation of AVs is essential. An online video-based training programme was developed to improve drivers' mental models for when the automation can be activated. A matched pairs experiment showed that this training programme in combination with an owner's manual was more effective in improving drivers' mental models for when the automation can be activated than reading an owner's manual in isolation. Additionally, drivers' trust in automation reduced after undergoing both types of training. This online training programme can support the safe use of AVs by helping drivers activate the automation in appropriate road conditions. This has the potential to create safer roads by reducing collisions linked to inappropriate mental models and trust in the automation. This way the clear benefits of AVs can be realised without the challenges.

Funding

This research was funded by IAM RoadSmart and the Engineering and Physical Sciences Research Council. These funders had no involvement in the study design, in the collection, analysis, and interpretation of the data, in the writing of the report, and in the decision to submit this article for publication.

CRedit authorship contribution statement

Siobhan E. Merriman: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Kirsten M.A. Revell:** Conceptualization, Methodology, Writing – review & editing, Funding acquisition. **Katherine L. Plant:** Conceptualization, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

References

- Allen, M. (2017). Matched groups. *The SAGE encyclopedia of communication research methods*.
- Bagloe, S. A., Tavana, M., Asadi, M., & Oliver, T. (2016). Autonomous vehicles: Challenges, opportunities, and future implications for transportation policies. *Journal of Modern Transportation*, 24, 284–303.
- Baldwin, T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, 41(1), 63–105.
- Barg-Walkow, L. H., & Rogers, W. A. (2016). The effect of incorrect reliability information on expectations, perceptions, and use of automation. *Human Factors*, 58(2), 242–260.
- Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, 18, 47–57.
- BMW. (2017, November 11). *The path of autonomous driving*. Retrieved March 30, 2020, from BMW: <https://www.bmw.com/en/automotive-life/autonomous-driving.html>.
- Boelhouver, 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.
- Cahour, B., & Forzy, J.-F. (2009). Does projection into use improve trust and exploration? An example with a cruise control system. *Safety Science*, 47(9), 1260–1270.
- Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692–702.
- Driver & Vehicle Standards Agency. (n.d.). *The official DVSA guide to hazard perception online*. Retrieved March 04, 2021, from The Official DVSA Learning Zone: <https://www.dvsalearningzone.co.uk/shop/hazard-perception>.
- DrivingTestSuccess. (2020, May 05). *Driving Theory test 4 in 1 kit app - what's included within this app?* Retrieved March 03, 2021, from <https://www.youtube.com/watch?v=RaCgChc960o>.
- Ebnali, M., Kian, C., Hulme, K., Ebnali-Heidari, M., & Mazloumi, A. (2019). User Experience in immersive VR-based serious game: an application in highly automated driving training. *Proceedings of the AHFE 2019 International Conference on Human Factors in Transportation*. 66, pp. 133–144. Washington D.C., USA: Springer.
- Endsley, M. R. (2017). Autonomous driving systems: a preliminary naturalistic study of the tesla model S. *Journal of Cognitive Engineering and Decision Making*, 11(3), 225–238.
- Fisher, D. L., Laurie, N. E., Glaser, R., Connerney, K., Pollatsek, A., Duffy, S. A., & Brock, J. (2002). Use of a fixed-base driving simulator to evaluate the effects of experience and PC-based risk awareness training on drivers' decisions. *Human Factors*, 44(2), 287–302.
- Ford Motor Company. (2016, August 16). *Ford of the Future: Autonomous Vehicles and Silicon Valley Investments | Innovation | Ford*. Retrieved May 13, 2020, from <https://www.youtube.com/watch?v=dDiHtI00hI>.
- Forster, Y., Hergeth, S., Naujoks, F., Krems, J., & Keinath, A. (2019). User education in automated driving: owner's manual and interactive tutorial support mental model formation and human-automation interaction. *Information*, 10(4), 1–22.
- Gentner, D., & Stevens, A. L. (1983). *Mental Models* (1 ed.). Psychology Press.
- Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in automation – Before and after the experience of take-over scenarios in a highly automated vehicle. *Procedia Manufacturing*, 3, 3025–3032.
- GOV.UK. (n.d.). *Theory test: cars*. Retrieved June 22, 2020, from Learn to drive a car: step by step: <https://www.gov.uk/theory-test/revision-and-practice?step-by-step-nav=e01e924b-9c7c-4c71-8241-66a575c2f61f>.
- Greenblatt, J. B., & Shaheen, S. (2015). Automated vehicles, on-demand mobility, and environmental impacts. *Current Sustainable/Renewable Energy Reports*, 2(3), 74–81.
- Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis. *Military Psychology*, 4(2), 63–74.
- Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. R. (2016). Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. *Human Factors*, 58(3), 509–519.
- Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. *Human Factors*, 59(3), 457–470.
- Hsu, C. C. (2012). Comparison of gender differences in young people's blog interface preferences and designs. *Displays*, 33(3), 119–128.
- Huang, J. L., & Ford, J. K. (2012). Driving locus of control and driving behaviors: Inducing change through driver training. *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(3), 358–368.
- Isler, R. B., Starkey, N. J., & Williamson, A. R. (2009). Video-based road commentary training improves hazard perception of young drivers in a dual task. *Accident Analysis and Prevention*, 41, 445–452.
- Isler, R. B., Starkey, N. J., & Sheppard, P. (2011). Effects of higher-order driving skill training on young, inexperienced drivers' on-road driving performance. *Accident Analysis and Prevention*, 43, 1818–1827.
- Korber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, 66, 18–31.
- Koustanai, A., Cavallo, V., Delhomme, P., & Mas, A. (2012). Simulator training with a forward collision warning system: Effects on driver-system interactions and driver trust. *Human Factors*, 54(5), 709–721.
- Krampell, M., Solís-Marcos, I., & Hjalmdahl, M. (2020). Driving automation state-of-mind: Using training to instigate rapid mental model development. *Applied Ergonomics*, 83, Article 102986.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.
- Lee, C., Ward, C., Raue, M., D'Ambrosio, L., & Coughlin, J.F. (2017). Age differences in acceptance of self-driving cars: a survey of perceptions and attitudes. In J. Zhou, & G. Salvendy, *Human Aspects of IT for the Aged Population. Aging, Design and User Experience. ITAP 2017. Lecture Notes in Computer Science* (pp. 3–13). Vancouver, Canada: Springer, Cham.
- Lexus. (n.d.). *OWNERS MANUAL*. Retrieved May 28, 2020, from Lexus: <https://drivers.lexus.com/t3Portal/document/omnav-s/OM48E92U/pdf/OM48E92U.pdf>.
- Manser, M. P., Noble, A. M., Machiani, S. G., Shortz, A., Klauer, S. G., Higgins, L., & Ahmadi, A. (2019). *Driver training research and guidelines for automated vehicle technology*. Texas, US: Safe-D National UTC, Texas A&M Transportation Institute, Virginia Tech Transportation Institute & San Diego State University.
- Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology*, 85, 971–986.
- Mehlenbacher, B., Wogalter, M. S., & Laughery, K. R. (2002). On the reading of product owner's manuals: Perceptions and product complexity. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, 46(6), 730–734.
- Mercedes-Benz Greenwich. (2019, February 13th). *What is Mercedes-Benz DISTRONIC PLUS®?* Retrieved April 27, 2020, from Mercedes-Benz: <https://www.mercedesbenzgreenwich.com/distronic-plus/>.
- Merriman, S. E., Plant, K. L., Revell, K. M., & Stanton, N. A. (2021a). Challenges for automated vehicle driver training: A thematic analysis from manual and automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 76, 238–268. <https://doi.org/10.1016/j.trf.2020.10.011>
- Merriman, S. E., Plant, K. L., Revell, K. M., & Stanton, N. A. (2021b). What can we learn from Automated Vehicle collisions? A deductive thematic analysis of five Automated Vehicle collisions. *Safety Science*, 141, Article 105320. <https://doi.org/10.1016/j.ssci.2021.105320>
- Merriman, S. E., Plant, K. L., Revell, K. M., & Stanton, N. A. (2023a). A new approach for training needs analysis: a case study using an automated vehicle. *Applied Ergonomics*, 111, Article 104014. <https://doi.org/10.1016/j.apergo.2023.104014>
- Merriman, S. E., Revell, K. M. A., & Plant, K. L. (2023b). Training for the safe activation of Automated Vehicles matters: Revealing the benefits of online training to creating glaringly better mental models and behaviour. *Applied Ergonomics*, 112, 104057. <https://doi.org/10.1016/j.apergo.2023.104057>
- Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors*, 38(2), 311–322.
- Montag, I., & Comrey, A. L. (1987). Internality and externality as correlates of involvement in fatal driving accidents. *Journal of Applied Psychology*, 72(3), 339–343.

- National Transportation Safety Board. (2020, February 25). *Collision between a sport utility vehicle operating with partial driving automation and a crash attenuator*. Retrieved April 23, 2020, from Accident Report Detail: <https://www.ntsb.gov/investigations/AccidentReports/Reports/HAR2001.pdf>.
- Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced “complacency”. *International Journal of Aviation Psychology*, 3(1), 1–23.
- Pollatsek, A., Narayanan, V., Pradhan, A., & Fisher, D. L. (2006). Using eye movements to evaluate a PC-based risk awareness and perception training program on a driving simulator. *Human Factors*, 48(3), 447–464.
- Pradhan, A. K., Pollatsek, A., Knodler, M., & Fisher, D. L. (2009). Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? *Ergonomics*, 52(6), 657–673.
- Revell, K. M., Richardson, J., Langdon, P., Bradley, M., Politis, I., Thompson, S., ... Stanton, N. A. (2020). Breaking the cycle of frustration: Applying Neisser's Perceptual Cycle Model to drivers of semi-autonomous vehicles. *Applied Ergonomics*, 85.
- Rudin-Brown, C. M., & Parker, H. A. (2004). Behavioural adaptation to adaptive cruise control (ACC): Implications for preventive strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(2), 59–76.
- Saffarian, M., de Winter, J. C., & Happee, R. (2012). Automated driving: human-factors issues and design solutions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 2296–2300.
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: a decade of progress. *Annual Review of Psychology*, 52, 471–499.
- Schoettle, B., & Sivak, M. (2014). *A survey of public opinion about autonomous and self-driving vehicles in the U.S., the U.K., and Australia*. Transportation Research Institute (UMTRI), Ann Arbor, Michigan. Retrieved from <https://deepblue.lib.umich.edu/bi>.
- ShareX. (n.d.). Retrieved March 20, 2021, from ShareX: <https://getsharex.com/>.
- Silva, P. A., Holden, K., & Nii, A. (2014). Smartphones, Smart Seniors, But Not-So-Smart Apps: A Heuristic Evaluation of Fitness Apps. In D. D. Schmorow, & C. M. Fidopiastis, Foundations of Augmented Cognition. Advancing Human Performance and Decision-Making through Adaptive Systems. AC 2014. (pp. 347–358). Springer, Cham.
- Society of Automotive Engineers International. (2018, June 15). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. Retrieved December 05, 2019, from SAE International: https://www.sae.org/standards/content/j3016_201806/.
- Sportillo, D., Paljic, A., & Ojeda, L. (2019). On-Road Evaluation of Autonomous Driving Training. *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 1-9). Daegu, South Korea: IEEE.
- Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science*, 47(2), 227–237.
- Stanton, N. A., & Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics*, 48(10), 1294–1313.
- Stanton, N. A., Walker, G. H., Young, M. S., Kazi, T., & Salmon, P. M. (2007). Changing drivers' minds: The evaluation of an advanced driver coaching system. *Ergonomics*, 50(8), 1209–1234.
- Stanton, N. A. (2015). Responses to autonomous vehicles. *Ingenia*, 62, 9.
- Sánchez-Franco, M. J. (2006). Exploring the influence of gender on the web usage via partial least squares. *Behaviour & Information Technology*, 25(1), 19–36.
- Taylor, T. G., Masserang, K. M., Pradhan, A. K., Divekar, G., Samuel, S., Butt, J. W., ... Fisher, D. L. (2011). Long term effects of hazard anticipation training on novice drivers measured on the open road. *Proceedings of the sixth international driving symposium on human factors in driver assessment, training, and vehicle design* (pp. 187-194). Olympic Valley - Lake Tahoe, California: PMC.
- Tengilimoglu, O., Carsten, O., & Wadud, Z. (2023). Implications of automated vehicles for physical road environment: A comprehensive review. *Transportation Research Part E: Logistics and Transportation Review*, 169, Article 102989.
- Tesla. (2019, May 16). *Model S Owner's Manual*. Retrieved May 17, 2021, from <https://tesla-info.com/doc/2019-tesla-model-s-eu.pdf>.
- TheoryTest.org.uk. (n.d.). *Free driving theory Test*. Retrieved March 03, 2021, from Theory Test: <https://theorytest.org.uk/free-driving-theory-test/>.
- Toyota. (2019, May). *Yaris Owner's Manual OM52A96*. Retrieved from <https://www.toyota.co.uk/>.
- Volvo Cars. (2017, April 03). *Volvo Cars CEO urges governments and car industry to share safety-related traffic data*. Retrieved May 13, 2020, from Volvo Cars: <https://www.media.volvocars.com/global/en-gb/media/pressreleases/207164/volvo-cars-ceo-urges-governments-and-car-industry-to-share-safety-related-traffic-data>.
- Walker, G. H., Stanton, N. A., Kazi, T. A., Salmon, P. M., & Jenkins, D. P. (2009). Does advanced driver training improve situation awareness? *Applied Ergonomics*, 40, 678–687.
- Watkins, I., Kules, B., Yuan, X., & Xie, B. (2014). Heuristic Evaluation of Healthy Eating Apps for Older Adults. *Journal of Consumer Health on the Internet*, 18(2), 105–127.