VeriCAV Report

Assessing Occupant Comfort

An Occupant Comfort Model for the evaluation of Automated Driving Systems (ADS)

June 2021

Martin Pett

Connected Places Catapult

VeriCAV

VERIFICATION OF CONNECTED AND AUTONOMOUS VEHICLES

Executive Summary

Background

VeriCAV is a modular and flexible platform to allow Automated Vehicles to be efficiently tested in simulated environments. VeriCAV automates the process of generating and analysing complex driving scenarios populated with realistically behaving road users.

VeriCAV integrates open-source, commercial and proprietary software systems, that together are capable of generating a multitude of road and traffic scenarios.

Autonomous Driving Systems (ADS) can be plugged into VeriCAV and tasked to navigate through thousands of scenarios in faster than real-time.

The VeriCAV platform analyses performance using a 'Test Oracle'. The Test Oracle observes and rates the performance ADS under test using algorithms. Each algorithm monitors a specific performance parameter, such as safety or road etiquette, as the ADS navigates through each scenario and assigns performance scores.

Developing the Comfort Algorithm

The aim of this VeriCAV sub-project, was to build a Comfort Algorithm that would sit within VeriCAV's Test Oracle and predict the occupant Comfort performance characteristics of an ADS.

Connected Places Catapult (also referred to as 'Connected Places') designed and built a comfort algorithm, based on scientific literature. Comfort parameters were selected to be objective and tailored to the context of Autonomous Vehicles. A scoring system was devised to rate the comfort performance of ADS within simulations which sat within the VeriCAV Test Oracle. The same scoring system was designed so that it could be used within the user trials to verify the comfort algorithm's accuracy and relevance to user trial participant's comfort experience.

User testing and verification

Connected Places Catapult created new software tools, processes and methods to duplicate VeriCAV scenarios and convert them into a format that participants could experience through Virtual Reality. Due to the 3rd Covid-19 Lockdown in 2020-21, the user trial could not be delivered using a driving simulator in the Connected Places Catapult's offices. Instead the user trials were delivered remotely using mobile VR technology and video conferencing, enabling participants to experience the simulated scenarios whilst working from home.

Outputs

The project delivered a Comfort Algorithm that measured and scored performance against five objective comfort parameters:

- Acceleration/Deceleration (Longitudinal and Lateral)
- Jerk (Longitudinal and Lateral)
- Headway (gap to vehicles directly ahead)
- Lateral Offset (proximity to adjacent vehicles)
- Gap acceptance (merging into traffic at T-Junctions and Roundabouts)

The comfort algorithm also produced an overall comfort rating for each scenario.

Verification

The user trial captured subjective and objective data that was analysed and compared to the comfort score generated by the comfort algorithm. Although there was good alignment between participant and algorithm comfort scores for Headway, Lateral Offset and Gap acceptance, there were notable differences for Acceleration and Jerk.

Using VR to expose participants to the scenarios proved to be successful for assessing and measuring comfort parameters that were based on visual/cognitive processing. However, the limited field of view and lack of peripheral vision resulted in poor perception of speed and acceleration. Although there are limitations of the VR system used, VR delivers a first person perspective experience of the simulation that is more immersive and experiential when compared to simulations presented on 2D monitors.

Conclusions

The tools and methodology in this project demonstrate great potential, not only to define how comfort could be assessed in a simulation, but it also demonstrates new ways in which humans/end users can be involved in the virtual testing process. There are subtle improvements to the comfort algorithm that will serve to align its test results to those of the participants and enable longer scenarios to be evaluated in the test oracle. The recommendations are to re-run the user trial once access to a driving simulator is possible, such that the acceleration and Jerk comfort parameters can be simulated and their impact on comfort can be appropriately evaluated.

Next Steps

CAVs have been promoted to deliver automated human transportation convenience, efficiency and improved road safety. However, across all road environments and conditions, we are still a distance from the transport utopia of a fully autonomous and connected transport system. There will be many difficult transition periods to navigate as technology evolves, as we progress through the SAE automation levels and manage the challenges of mixed vehicle fleets on our roads.

In addition to these technological transition challenges, gaining public trust in autonomous systems is critical to achieving rapid and widescale adoption. The industry needs to make efforts to increase end user involvement in the development of their products and services to ensure there is acceptance, trust and demand.

This project demonstrates that there are methods and technologies available today that will support and bridge the gap between what is technically achievable and what humans will accept, want and embrace. Involving end users in the development process, especially early on, will deliver considerable benefits due to faster and larger uptake of products and services offered to the public.

Connected Places Catapult welcomes interested parties to get in contact to discuss this project and explore collaborative partnerships to evolve the technology for commercial use.

Martin Pett

Comfort Model and User Trial Project Lead

martin.pett@cp.catapult.org.uk

Table of Contents

Executive Summary 2
Background 2
Developing the Comfort Algorithm 2
User testing and verification 2
Outputs 2
Verification 3
Conclusions 3
Next Steps 3
Notice 6
Authorisation: 6
Record of changes: 6
1 Introduction 7
1.1 VeriCAV project overview 7
1.2 The VeriCAV comfort work package 7
2 Research and Development of the VeriCAV Comfort Algorithm 9
2.1 The Comfort Algorithm 9
2.2 Comfort Factors 11
2.3 Comfort Thresholds 12
2.4 Acceleration and Jerk 12
2.5 Lateral offset 13
2.6 Headway 14
2.7 Turnings, junctions and roundabouts 14
2.8 Comfort scoring system 15
3 Designing the User trials 17
3.1 Objective 17
3.2 Approach: Building the simulation and delivering the simulator 17
3.3 Approach: Designing the User trial/experiment 18
4 Method 20
4.1 Approach 20
4.2 Technical Approach 20
4.3 Software development to support the user trial 21
4.4 Pivoting from the original technical approach (COVID-19) 21
4.5 Decontamination 23
4.6 Experimental Method 24
5 Results 27

5.1 Familiarisation and Speed Calibration 27
6 Comfort Results 30
6.1 Overall Scenario Rating Summary: 30
6.2 SCENARIO COMFORT DATA ANALYSIS 31
6.3 Scenario 1: Motorway merge 31
6.4 Scenario 2: Motorway lane change and Exit 32
6.5 Scenario 3: Roundabout no traffic 33
6.6 Scenario 4: Roundabout with traffic 34
6.7 Scenario 5: A Road oncoming car in adjacent lane 35
6.8 Scenario 6: A Road, slowing down for slow car ahead. 36
6.9 Scenario 7: T Junction no traffic 37
6.10 Scenario 8: T Junction with traffic 38
7 Discussion 39
7.1 Comfort Factors 39
7.2 Comfort Model review 43
7.3 Participants 44
7.4 Experimental method 44
7.5 Ego Vehicle / Other Vehicle Behaviour 45
7.6 Statistics and data visualisation 46
8 Conclusion 48
8.1 Comfort Algorithm vs. Participant Comfort Ratings 48
8.2 Overall Comfort Rating 48
8.3 Summary 49
9 Limitations 51
9.1 Technical 51
9.2 Methods 52
10 Next Steps 54
10.1 Refine comfort model 54
10.2 Rerun the user trial simulation with motion platform 54
10.3 Analyse the data with University partners 54
10.4 Was using VR the right choice? 54
10.5 The Future of VR user trials? 55
10.6 Final Remarks 56
11 Outcomes and Innovation 56
References 57
Bibliography 57
Acknowledgements 60

Appendix A 61
Comfort Data 61
Appendix B 89
Participant demographics 89

Notice

This document has 94 pages including the cover.

Authorisation:

Action	Name	Position in Company	Name	Position in Consortium
Written by:	**Martin Pett**	**Principal Technologist**	**Shyma Jundi**	**Senior Technologist**
Reviewed by:	**Alan Peters**	**Solutions Architect**	**Jamie Chan-Pensley**	**Principal Technologist**
Authorised by:	**Andrea Cooper**	**Director of HCD / CPC**	**Henry Tse**	**Director of HCD / CPC**

Record of changes:

Released to	Version	Reason for Change	Date

1 Introduction

1.1 VeriCAV project overview

VeriCAV (Verification of Connected and Autonomous Vehicles) has been a multi-million pound, 27 month collaborative research project involving a partnership of four organisations: HORIBA MIRA (as industry lead), Connected Places Catapult, the University of Leeds, and Aimsun. The goal was to create a framework to allow efficient testing of Automated Driving Systems (ADS) in simulation.

Figure 1: Shows the VeriCAV framework to enable the efficient testing of Automated Driving Systems (ADS)

Automated vehicles will transform the way we use transport, unlocking time currently spent driving, reducing accidents from human error, and improving mobility for those who can't easily drive themselves.

To achieve this vision, automated vehicles need to be tested robustly to ensure their safety before deployment. The complexity and variability of driving in a public road environment makes overcoming this barrier a significant challenge. One increasingly important technique involves testing a range of scenarios in simulation.

Testing in simulation allows developers to have full control, enabling manipulation of test conditions, ease of testing within a safe environment, with the ability to run and repeat large numbers of tests in parallel. All of these properties leading to an increase in reliability and validity of the results. The challenge is ensuring realism of the simulation whilst avoiding the need to formulate and evaluate each test scenario manually. This is the challenge VeriCAV addresses.

1.2 The VeriCAV comfort work package

The objective of the comfort work package was to develop a comfort algorithm that VeriCAV's Test Oracle could use to determine/predict passenger comfort.

To achieve this, the comfort work package developed an algorithm, based on scientific literature and a scoring system that the VeriCAV test oracle would use to rate the performance of an ADS. The comfort algorithm was one of the test algorithms developed in the course of the VeriCAV programme.

To verify that the comfort algorithm was representative of the comfort experienced by end users, the work package designed and ran a user trial to collect objective data and subjective feedback.

Virtual Reality was selected as the most effective and efficient method to expose participants to simulated scenarios and collect data. VR headsets were pre-loaded with scenarios and couriered to participants who were working from home due to the 3rd national lockdown during the Covid-19 Pandemic of 2020-21.

Using pre-recorded video tutorials and online video conferencing, the Research and Human Factors team were able guide participants how to use the VR equipment and navigate them through the user trial.

2 Research and Development of the VeriCAV Comfort Algorithm

2.1 The Comfort Algorithm

An algorithm was developed based on a scientific literature review, which combined the data from a range of research papers. Comfort is a highly subjective, personal and contextual concept which has led to varying definitions of comfort.

There are many ways that scientists and researchers have defined and grouped the factors that contribute to comfort with divergent views on the grouping and hierarchy. Some definitions are focussed on the ergonomic aspects, combined with experience “comfort is the result of correct ergonomic design and additional convenience features” Siebertz 2014. Siebertz states that:

1. All comfort aspects must be considered simultaneously.
2. A comfortable vehicle has a good performance in every comfort discipline.
3. Designing for comfort requires deep knowledge of the human physiology.

Error! Reference source not found. below shows a conceptual model that is predominantly based on the cognitive/psychological factors of comfort.

Figure 2:Shows the conceptual framework for the design of the passenger experience in Shared Automated Vehicles (SAV) "Designing for Comfort in Shared and Automated Vehicles (SAV): a Conceptual Framework (Cyriel Diels et al. 2017)"

Although there was recognition of the importance of the many cognitive comfort factors, there was an acceptance that defining a comfort algorithm to measure cognitive comfort would extremely challenging in the scope of this project due to the multiple facets, the continuously fluctuating and largely personal nature of comfort as a construct.

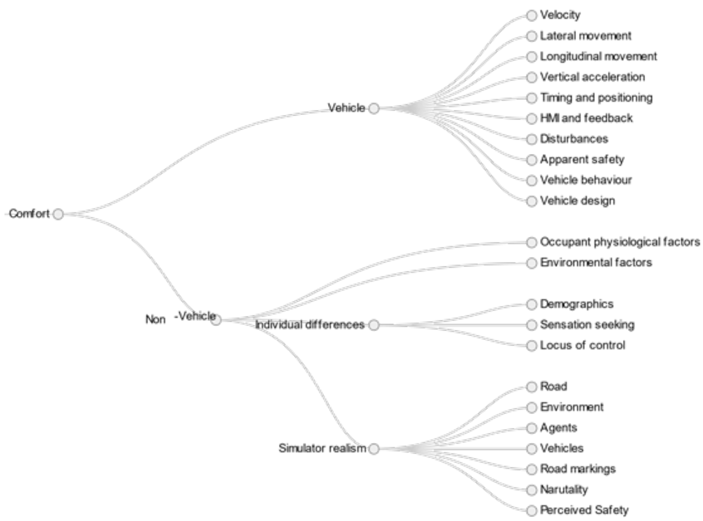


Figure 3: Shows a subset of the comfort factors identified in the research that could be incorporated into an algorithm and constrained in a user trial using a simulation.

The researchers mapped the comfort factors identified in the literature research. A subset of these is shown in Figure 3 above. The comfort factors were ordered in terms of comfort hierarchy, importance and what was feasible to test given the project constraints. The constraints were:

- What could be objectively included in an algorithm
- the working environment of potential user trial participants.

2.2 Comfort Factors

The following comfort factors were selected for the comfort algorithm, acknowledging that the algorithm would need further development in the future to refine and potentially add more parameters to fully reflect the complexity of measuring occupant comfort in Autonomous Vehicles:

1. Acceleration (longitudinal and lateral) - Definition of acceleration
2. Jerk (longitudinal and lateral) - Definition of Jerk
3. Headway to vehicles ahead (in the same lane) – Definition of Headway
4. Proximity to vehicles in an adjacent lane (oncoming and overtaking)
5. Gaps between vehicles when joining a new carriageway (T-Junction and roundabouts)

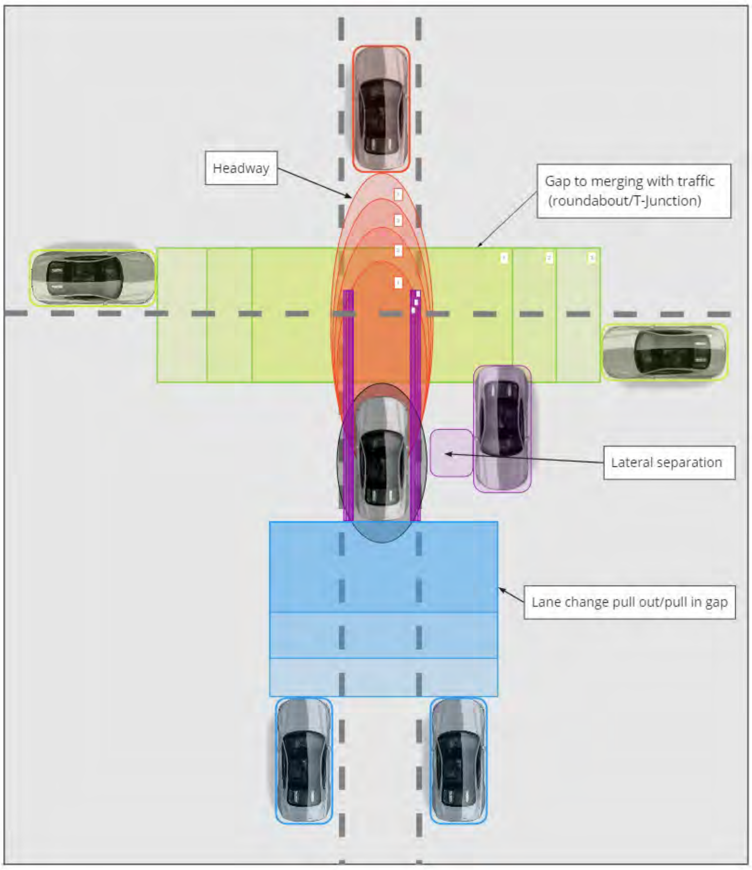


Figure 4: A graphical representation of the comfort factors and thresholds included in this study

2.3 Comfort Thresholds

The comfort algorithm replicated the literature's terminology for 'levels' of comfort and the objective thresholds/boundaries for each comfort level.

The comfort levels were:

- Comfortable
- Normal
- Uncomfortable/ Aggressive
- Very Uncomfortable/Very Aggressive

The following sections identifies how the comfort threshold levels identified in the theory relates to the different factors of the algorithm.

2.4 Acceleration and Jerk

The table below shows the comfort levels and their objective boundaries for Acceleration & Jerk:

Table 1 - Chart showing accelerative comfort factor thresholds

Accelerative Comfort Factors								
	Acceleration m/s2, Jerk m/s3							
Parameter	**Comfort**		**Normal**		**Aggressive**		**Extreme Aggressive**	
ACCELERATION								
Lateral Acceleration	-0.9	0.9	-4	4	-5.6	5.6	-7.6	7.6
Longitudinal Acceleration	-0.9	0.9	-2	1.47	-5.08	3.07	-7.6	7.6
JERK								
Lateral Jerk Acceleration	-0.6	0.6	-0.9	0.9	-2	2	>-2	>2
Longitudinal Jerk Acceleration	-0.6	0.6	-0.9	0.9	-2	2	>-2	>2

The literature showed that acceleration in longitudinal and lateral directions should not be assessed in isolation. For example, if a car was accelerating hard in a straight line, that might be acceptable given an appropriate circumstance, however, if the car was accelerating hard whilst going round a tight bend, then this would be uncomfortable and the threshold boundaries adjusted due to the cumulative effect. It was therefore decided to adopt the approach proposed in paper by Bae, J. Moon and J. Seo 2019, whereby the boundary thresholds for lateral and longitudinal acceleration and jerk were combined, creating diamond profile boundaries. See Figure 5 below:

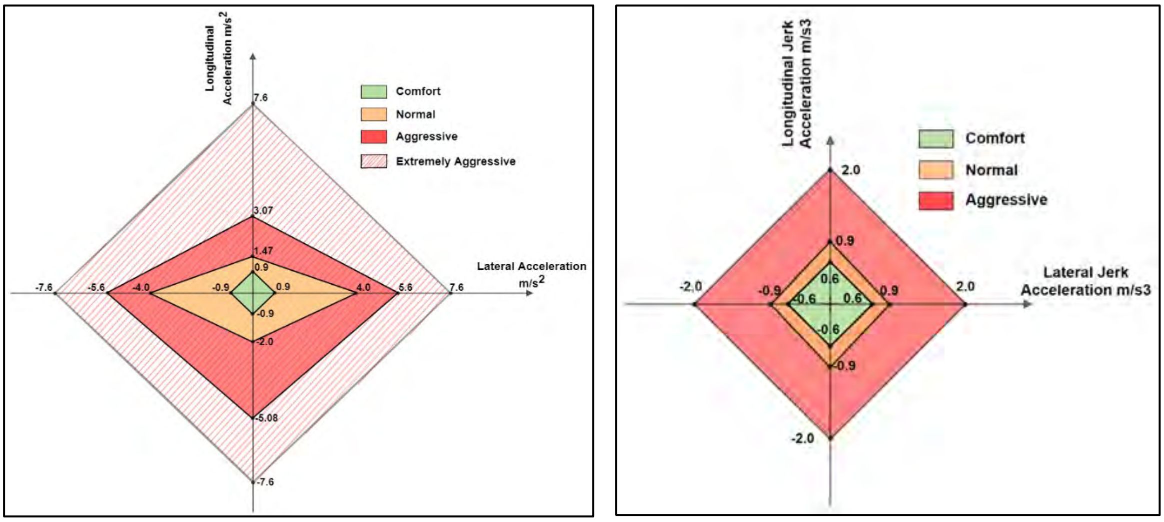


Figure 5: Shows the boundaries defined for Acceleration and Jerk (longitudinal and lateral)

2.5 Lateral offset

The comfort boundaries and comfort levels for Lateral offset/proximity to adjacent vehicles were defined as follows:

Table 2- Shows the comfort categories for lateral offset to adjacent vehicles

Lateral Offset distance from oncoming & adjacent vehicles	
Separation in Metres to oncoming/ passing vehicles	**Comfort Level**
>0.8	**Comfortable**
0.68 - 0.8	**Normal**
0.43 - 0.68	**Aggressive/ Uncomfortable**
<0.43	**Extremely Aggressive/ Very Uncomfortable**

The literature, Rossner and Bullinger, 2019 and Gudrum et al. study 2018, varied in their methodology and evaluation techniques to measure lateral offset comfort. However, from the literature, it was possible to quantify the overall distance between vehicles (side of vehicle to adjacent vehicle), regardless of lane width and positioning, and participant's levels of comfort. The values shown in Table 2, should be considered with the caveat that lane width, vehicle size, weather and road conditions, all have an effect on user acceptance of lateral offset, for which the table does not accommodate/represent. Further research and refinements are necessary to take into consideration these additional factors.

2.6 Headway

The comfort levels and boundaries for Headway were defined as follows:

Table 3 - Shows the relationship to time separation to vehicles ahead and comfort.

Headway to car ahead			
Time (Seconds)	Speed		
	50km/h	100km/h	150km/h
>4.0	Comfortable	Comfortable	Comfortable
2.0 - 4.0	Comfortable	Comfortable	Comfortable
1.5 - 2.0	Normal driving	Comfortable	Normal driving
1.0 - 1.5	Aggressive driving	Normal driving	Aggressive driving
0.5 - 1.0	Extremely aggressive	Extremely aggressive	Extremely aggressive
0.0 - 0.5	Extremely aggressive	Extremely aggressive	Extremely aggressive

The literature identified for headway aligned well to the comfort level terminology. There were some interesting relationships between speed and headway gap (seconds). For more information about the above table, refer to the paper by Felix Wilhelm Siebert et al. 2014.

2.7 Turnings, junctions and roundabouts

The comfort levels and boundaries for pulling out of junctions were as follows:

Table 4: Shows the gap between vehicles and the associate comfort categories.

Turning and joining lane (T-Junction) & (Roundabouts)	
Ave gap (Seconds) between vehicles	Comfort Level
10	Comfortable
8	Normal
6	Aggressive
4	Extremely Aggressive

The paper Felix Wilhelm Siebert et al. 2014, from which the above data is taken, used CCTV to measure the gap between vehicles when cars pulled out from a T-junction. This was based on observations of 10,419 drivers and 22,630 gaps observed. The decision was made to use the same gaps between vehicles at T-Junctions and apply them to roundabout scenarios. (i.e. the distance/time between the Ego and vehicles on the roundabout when the Ego joins a roundabout).

2.8 Comfort scoring system

Having defined the comfort factors, terminology and the objective thresholds for comfort, the next step was to define a method to enable the Test Oracle to measure the comfort performance of an ADS.

To achieve this a scoring system was designed that was capable of rating the performance of the individual comfort factors as well as an overall (Scenario) comfort score. It was deemed useful to present the score/performance of each comfort factor over time, as this would help engineers developing ADS' to identify the specific behaviour or interaction that resulted in negative scoring.

The scoring matrix was devised to score the ADS against each comfort factor over 10 second time segments. 10 second segments were chosen to provide enough granularity, such that engineers could identify where the uncomfortable behaviour occurred within a scenario. Furthermore, 10 second segments would allow for longer scenario lengths in the future.

Each comfort factor starts with 100 points in each 10 second segment. Deductions are made when the ADS/data exceeds comfort boundaries. i.e. When a comfort factor's data moves from 'Comfortable' into 'Normal', Aggressive or Very Aggressive. The scoring is incrementally more punitive, relative to the comfort boundaries as shown in the Scoring Matrix below:

Table 5: Shows the negative scores that will be deducted from 100points per second.

Scoring matrix		
	Sample time **10s**	**Starting Score** **100pts**
Comfort levels	**<u>Negative</u>** Scoring / sec	Secs of exposure to result in a Fail
Comfortable	0	n/a
Normal	-1	40
Aggressive/ Uncomfortable	-10	4
Extremely Aggressive/ very uncomfortable	-20	2

To grade the performance of each comfort factor and the overall performance of an ADS in a scenario, the following comfort grading matrix is used.

Table 6: Shows the comfort Scoring matrix that is used to grade the individual comfort factors AND the overall performance of an ADS in a scenario.

Comfort Grading Matrix		
Comfort Grades	**Comfort Score**	**Description**
A*	>90	Extremely comfortable
A	>80	Very comfortable
B	>75	comfortable
C	>65	Acceptable
D	>60	Poor
F	<60	Fail / uncomfortable

The Comfort Grading Matrix is applied to the scores of each comfort factor over 10 second sample times. If any comfort factor scores less than 60 in any 10 second time sample, then the ADS fails the comfort test. Otherwise the grades are applied for every 10 second sample time and the average score is graded accordingly.

The same Comfort Grading Matrix is also used to provide an overall comfort grade for the ADS per scenario. To do this, the average score for each comfort factor are added together and then averaged. This shows the overall performance for the ADS. If an 'F' grade is awarded anywhere in the scoring, the ADS will have deemed to have Failed the comfort test. This approach was based on the assumption that a significant negative comfort experience would affect the overall comfort rating of a scenario making the experience unacceptable to occupants/participants.

3 Designing the User trials

3.1 Objective

The user trials were designed to:

- Compare the comfort ratings of participants to those generated by the comfort algorithm.

This would be achieved by performing the following:

- Recording participant perceptions for each comfort factor
- Recording participants' overall comfort rating for each scenario
- Identify instances within a scenario where they felt uncomfortable
- Capture verbatims to add situational context to when, where and why they felt uncomfortable.

3.2 Approach: Building the simulation and delivering the simulator

To achieve the objectives, it was imperative that participants experienced simulated scenarios (generated by the VeriCAV simulation software), in such a way that there was consistency. i.e. How the ADS drives through a simulated environment in the VeriCAV simulation software matches exactly what participants experience in the VR simulation.

There was a challenge in applying this in theory because what the ADS 'sees' in the VeriCAV simulation software is different to what Humans will expect/experience in a simulated environment. For example, the ADS 'sees' the road by reading the OpenDrive and OpenScenario files and manoeuvres accordingly. For the user trial, the OpenDrive files needed to be crafted into a realistic 3D virtual environment, with realistic vehicles that moved naturally through the scenario.

Other differences exist between a simulation and a simulator. For example, a computer can calculate the acceleration and know from the physics/calculations that acceleration is X. However, in a simulator, it is very hard for a participant to determine the speed or acceleration because they cannot physically experience/sense the G-Forces that they instinctively associate with acceleration. If participants cannot feel acceleration, can they accurately assess acceleration in a simulator? This is a known challenge in the simulator industry and there are numerous ways to manage this divergence between reality and simulation. Motion platforms have been used to recreate motion that replicates the effects of G-force, by tilting and jolting the platform where users are sat.

Another challenge is the 'naturality' of the movement in a simulation versus the reality of what happens in the real world. This is called deterministic and non-deterministic simulation. The former can be described as an 'on rails' experience. The simulated vehicles go exactly where they are programmed to go. The latter is a simulation whereby the 'Ego' vehicle (the one in which the participant experiences the simulation) and the other 'agents' in the simulation are programmed to move from A to B with some parameters and constraints to adhere to. So for example, if the vehicle is programmed to go from A to B round a curve as fast as possible and the road is wet, then the simulation will attempt to achieve the goal, but there may be some traction loss that could result in the car steering/braking to achieve the end goal. The end goal is the same, but the experience is more natural and lends itself to a simulator that humans can experience with greater levels of perceived realism.

With acknowledgement of the limitations of the experimental method, the researchers accepted these challenges and planned strategies and methods that would limit influence on the results.

This meant that the simulated environment should not be too detailed or too high resolution, as this may encourage the participant to look out the car's side windows, thus distracting them from the interaction/focal point of the scenario. In addition, the decision was made to repeat environmental features where appropriate with only minor subtle changes throughout a scenario. For example, one type of tree, bush, grass, road texture and only three building types were used to populate the scenarios. This also had the benefit of reducing computational demands, which are always a challenge when using VR.

To account for the physics of Acceleration and Jerk, the plan was to use the Connected Places Catapult's driving simulator with a wide-angle field of view VR headset. The Connected Places driving simulator is fitted with a motion platform system made by D-Box. This is a motion system used extensively by the car racing industry. If the D-Box system could provide motion cues to simulate acceleration and jerk, then it was deemed that this would be a good feature to include.

The decision to use of a wide-angle field of view VR headset was driven by the literature that concludes that a significant part of a human's ability to detect and asses speed / acceleration, uses peripheral vision. A wide-angle VR headset was therefore purchased for the user trial.

Figure 6:Shows the Connected Places Catapult's driving simulator with D-Box motion platform and the Pimax 8K plus VR headset that has a 210° Field of View.

3.3 Approach: Designing the User trial/experiment

Data surrounding other comfort factors identified in the literature review was simultaneously captured throughout the trial activity.

This data was collected via questionnaires administered prior to and post the VR user trial. The following diagram shows how the user trial was designed to mirror the comfort factors of the Comfort Algorithm and the other research topics that were considered for inclusion in pre and post-trial questionnaires.

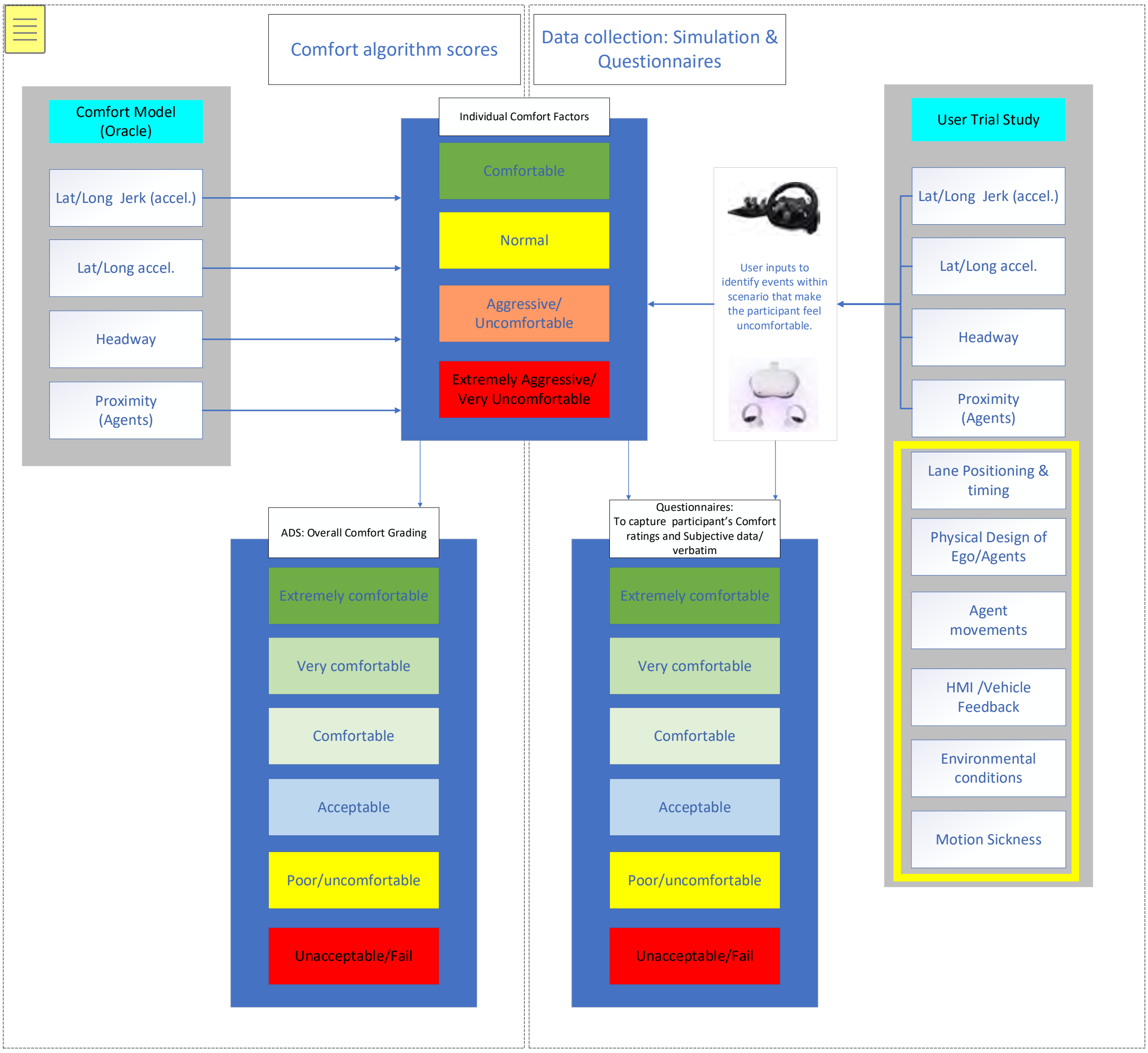


Figure 7: The diagram shows how the user trial was designed to collect data that mirrored the comfort algorithm's comfort ratings. The boxes highlighted in bright yellow shows the extended research questions that the Connected Places Catapult's researchers aimed to investigate through questionnaires pre and post-trial.

4 Method

4.1 Approach

The primary goal of the user trial was to design a study that would enable direct comparison of the comfort score generated by the comfort algorithm and the results from the user trial.

To achieve this, the scenarios and the scoring metrics were identical. This meant that the scenarios, timing and movements tested by the comfort algorithm had to be identical to those used and experienced in the user trial.

4.2 Technical Approach

Using OpenDrive and OpenScenario standards, eight scenarios were created to provide a range of environments, road types and challenges for an ADS to navigate.

The eight scenarios were designed to have a duration of 30-70 seconds. This was slightly longer than the normal scenarios generated by VeriCAV. Adding 5 seconds on to the beginning or end of each scenario gave participants time to gain contextual awareness of their surroundings before encountering a road junction and/or other vehicles. See Figure 8 for an example of the storyboards created from the OpenDrive maps and how they were enhanced to support the VR simulation/user trial.

The eight scenarios created were:

- Scenario 1: Merging onto a Motorway
- Scenario 2: Changing lane on a motorway and taking a motorway exit slip road
- Scenario 3: Joining a roundabout and exiting (no cars)
- Scenario 4: (***Apollo***): Joining a roundabout and exiting (negotiating 2 cars on the roundabout)
- Scenario 5: Driving on an A road and passing an oncoming car
- Scenario 6: Driving on an A road, following a car ahead and then slowing down due to a slower car ahead
- Scenario 7: Turning right at a T-Junction (no other cars)
- Scenario 8 (***Apollo***): Turning right at a T-Junction (with cars approaching the T-junction in both directions)

The scenarios were input into the VeriCAV software suite and a log file was generated for each scenario. This log file recorded the position and movements of the Ego and other vehicles time steps throughout each scenario.

> ***Note:*** *Apollo was used to drive the movements of the Ego in Scenarios 4 and 8. For the other scenarios OpenScenario was used to co-ordinate the movements of the Ego and other vehicles in the scenarios.*

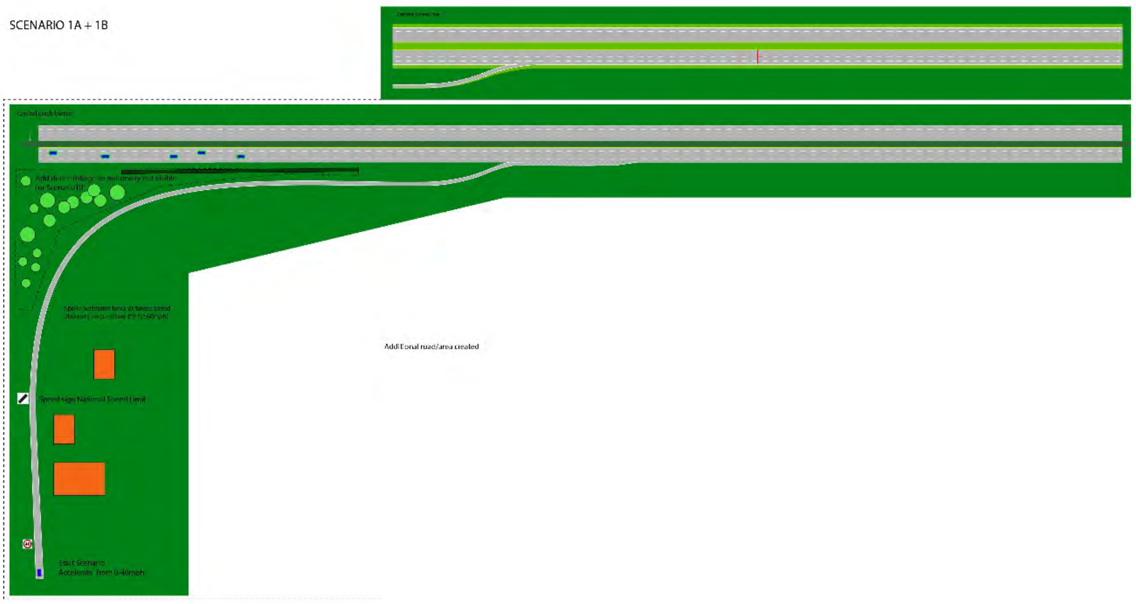


Figure 8: Shows the original open drive map (top), the enhanced version below that includes additional road to support participant acclimatisation to the scenario and the buildings, signage and foliage.

The log files for each scenario were then processed through the VeriCAV Test Oracle to generate scores for each comfort factor, an overall comfort rating and a test PASS or FAIL result.

4.3 Software development to support the user trial

The Connected Places software development team used the files generated for and outputted by the VeriCAV software suite/Test Oracle. The Team added features, functionality and 3D environments to create the necessary capabilities to deliver the VR user trial.

This included adding materials and correcting/refining the geometry, adding line markings, road signage, buildings, foliage, and vehicles to each scenario. The Ego vehicle did not have an interior, so an interior was added to the exterior 3D model to create a realistic and dynamic interior where participants would be sat (in VR).

The Ego's instrument display's indicator lights were animated, and an auditory tick tock was programmed to tick in sequence with the indicator lights. Road noise was added to create a more realistic driving experience. Vehicles in the scenario applied their brake lights and indicators in the scenarios as appropriate. Materials and lighting were also configured to create a realistic experience.

4.4 Pivoting from the original technical approach (COVID-19)

Due to Covid-19 restrictions, the team pivoted from an office-based user trial and developed a bespoke software solution that enabled participants to experience the VR simulation/scenario from the comfort of their home's office.

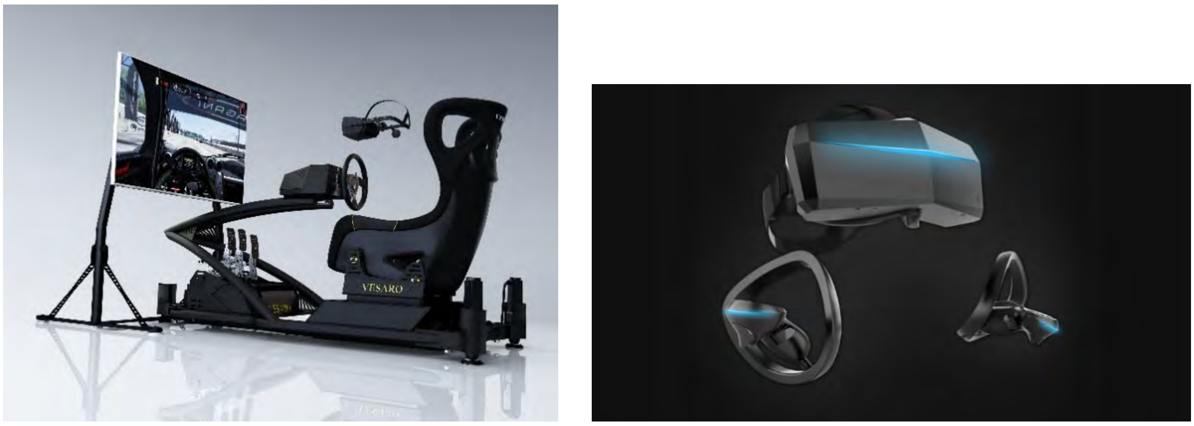


Figure 9: Shows the original hardware plans for delivering a VR experience.

This novel approach meant that the simulation could no longer rely on a High-Powered gaming PC, but instead had to be compatible with a significantly lower computing power mobile VR headset.

The simulations/scenarios created in the Unity Gaming Engine and LGSVL software were recorded within Unity by a 360 video capture software package thus capturing a 1st person perspective that could be replayed in the VR headset. This approach removed the need for high computing power/processing associated with running the simulation in real-time.

Figure 10: Shows the Oculus Quest 2 Enterprise VR headsets that were dispatched to participant's homes for the user trial and a remote user trial delivered using MS Teams.

A dedicated VeriCAV application (Figure 11), was built with a menu interface that enabled users to select tutorial videos, familiarisation scenarios and the eight user trial scenarios within the VR headset.

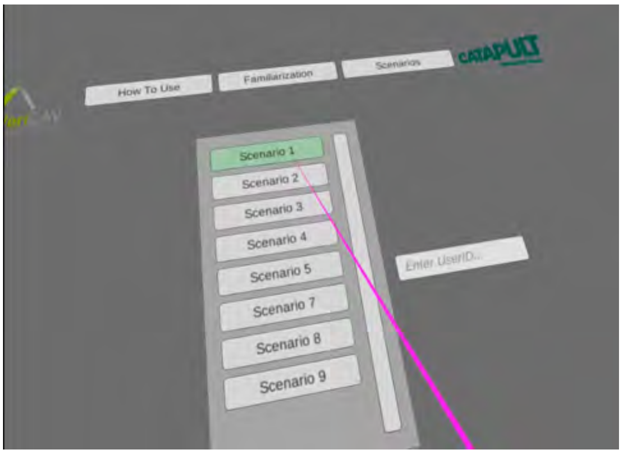


Figure 11: Shows the VeriCAV VR menu system.

The VeriCAV user trial application was also designed to record the following:

- Log the participant's unique ID numbers,
- Head movement
- Button presses (via the hand-held controllers)

A further desktop application was developed that would use the data captured on the headset and enable researchers to review participant's VR experiences. The purpose of this software was to help researchers identify:

When and where in the scenario the participant felt uncomfortable (utilising the recorded button presses)

Where they were looking (utilising the monitored head movement and orientation).

This was understood to be a novel approach to running user trials and a unique approach for user experience of AVs by capturing data from participants who are situated in different geographical locations. Although the ability for participants to detect acceleration and jerk was expected to be compromised (no motion platform), the potential this approach demonstrated for distributed large-scale user trials looks to be very promising and has significant potential if applied correctly.

4.5 Decontamination

Due to the physical contact of the VR headset with participants, there was a need to perform a thorough cleaning activity prior to each participant using the VR headset. This was as follows:

1. VR headset was sent out/returned to the office before/after each participant.
2. The surface of the VR headset and controllers were cleaned with an antibacterial wipe.
3. The headset was placed into a CLEANBOX UV sterilisation system for decontamination.
4. The headset was removed from the CLEANBOX system using protective gloves and placed back into the transportation box.
5. Antibacterial/virus wipes and gloves were included with the VR headset for participants to wipe down upon receipt if they feel this was necessary.
6. Participants were asked to use provided gloves and antibacterial/virus wipes to clean the headset after they completed the user trial and prior to putting it back into the box.

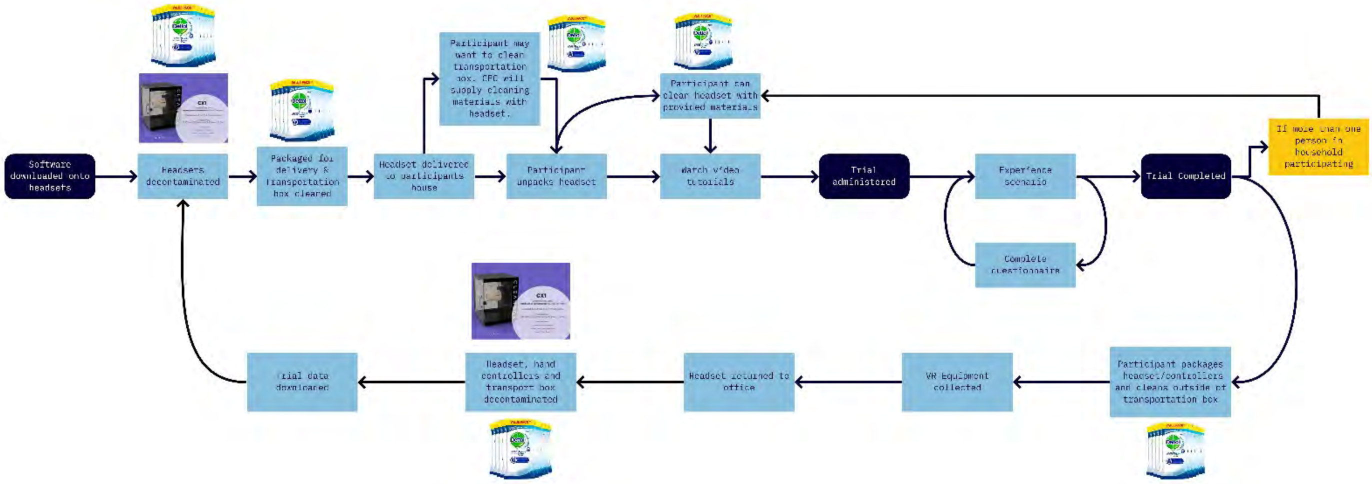


Figure 12: Shows the process used to ensure the VR Headset was thoroughly decontaminated throughout the user trials

4.6 Experimental Method

The user trials were split into three sections:

1. Pre-trial – participant recruitment and selection
2. User trials – Real-time user trials
3. Post-trial – Post trial feedback

4.6.1 Pre-trial

Pre-trials included recruitment of participants through an email to Connected Places Catapult staff and family members around the Buckinghamshire and Northamptonshire areas. This was performed to control the distribution of VR headsets and to enable a greater level of control on the safety and well-being of the participants throughout the trial.

Participants were selected from the pool of questionnaire respondents and a spread of males and females and a range of ages were invited to participate. The participants were also screened for their susceptibility to motion sickness, their approximate annual mileage and a brief questionnaire on their technological orientation.

Participant recruitment questionnaire:

- Demographic information
- Motion Discomfort questions
- Attitudes towards new technologies (ATTTECH)
- Attitudes towards automation (ATTAUTO)
- Attitudes towards CAVs (ATTCAV)

Once participants were notified of the trial, they were provided with a pre-trial questionnaire which consisted of a set of different psychometric tests pertaining to their driving behaviour, their interest in participating in exciting activities and their belief of being in control of events (locus of control).

Pre-trial questionnaire:

- Driver Behaviour Questionnaire (DBQ)
- Traffic Locus of Control (TLOC)
- Arnett Inventory of Sensation Seeking (AISS)

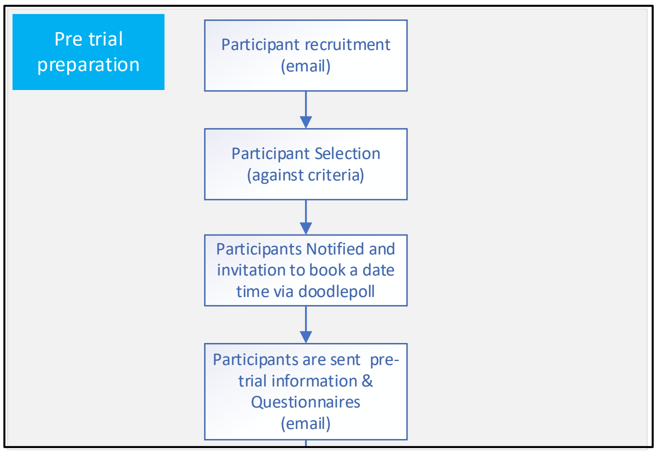


Figure 13- Pre-trial procedure

4.6.2 User Trials

The User trials were conducted using a combination of a Virtual Reality (VR) headset – an Oculus Quest 2 to simulate various in-vehicle autonomous vehicle scenarios, a video conferencing facility – Microsoft Teams, to conduct real-time interactions with the participant and to monitor their actions and progress and an online questionnaire application to capture their feedback.

Firstly, the participant was welcomed to the trial. They were then familiarised with the equipment (if they hadn't done so beforehand) and were then asked to perform the first of the three familiarisation activities. After each familiarisation session, the participant was asked to identify what speed that they thought the simulated vehicle was travelling at. The participant was not provided with a working speedometer or any cues to determine what the speed was, it was entirely down to their own perception. This was input into the 'in-trial' questionnaire. The familiarisation process and results can be found in the Familiarisation and Speed Calibration section.

After the familiarisation session, the participants were exposed to eight scenarios one by one. After each scenario was shown, the participant was asked a series of questions that identified any points in the scenario where the participants felt uncomfortable (and pressed the controller's buttons) or by verbal expressions/comments. The scenarios during the trial and the familiarisation session were randomised to ensure there is no order bias, with the facilitator instructing the participant on the order of the scenarios they were to open.

Responses to the structured questions and verbatim were recorded by the facilitator in the in-trial questionnaire form.

Participants were also checked for motion sickness after each scenario to make sure that they were not experiencing any negative effects from the VR system and given the option to take a break. Before the start of each scenario, the participants will be asked to check the headset is fitted correctly and is in focus, to ensure that they are comfortable and could clearly see the videos presented.

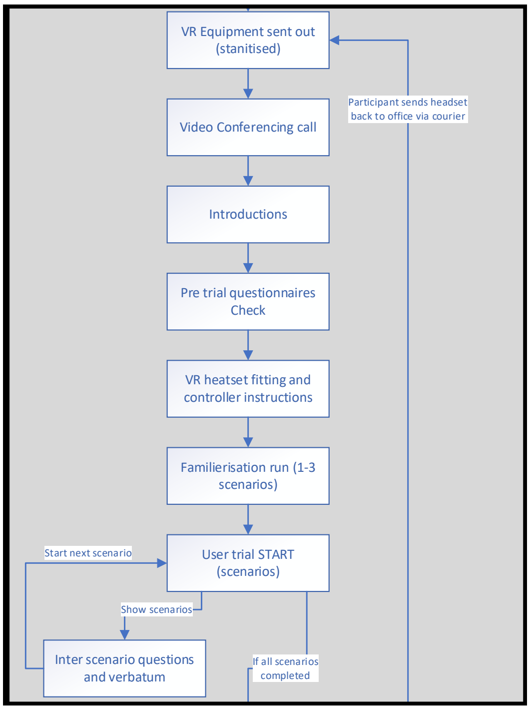


Figure 14: Show the in-trial process

4.6.3 In-trial questionnaire:

In the trial, after each scenario had been run, the participants were directed to answer the in-trial questionnaire. This contained a series of standardised questions. This was administered using Survey Monkey, an online questionnaire software tool.

4.6.4 Post Trial

After each participant had experienced all eight scenarios, they were led through a series of post-trial questions. These focussed on the simulation quality, realism and ease of use of the equipment. They were also asked to comment on their feelings of motion sickness and also provided a section where they were able to comment on any improvements to the in-vehicle user interface and how the user experience within an autonomous vehicle may be improved.

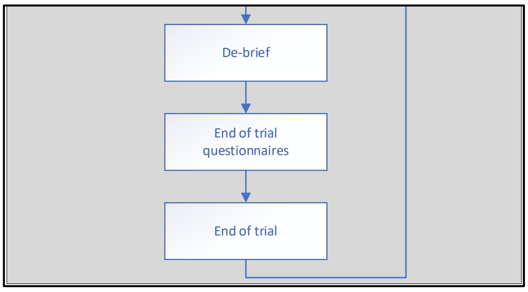


Figure 15: Shows the post trials process

4.6.5 Post-trial questionnaire:

- Simulator Realism Questionnaire (SIMR)
- Human Machine Interface (HMI) questionnaire

5 Results

5.1 Familiarisation and Speed Calibration

A familiarisation process was conducted for participants at the beginning of each study. This was to ensure that participants felt comfortable with the headset, the surroundings in which they would experience each scenario and the format of each exposure.

The aim of this process was also to gain a baseline understanding of the participants perception of speed in the simulation. Each participant was exposed to three conditions in a randomised order, these were as follows:

1. 40 miles per hour
2. 50 miles per hour
3. 60 miles per hour

Each exposure was 26 seconds in duration, which was slightly shorter than the experimental scenarios, but was deemed as sufficient for participants to draw a judgement on the speed.

These familiarisation sessions consisted of a small curve leading to a straight road, populated with a medium density of trees, hedges and buildings, with a vehicle passing in the opposite direction. Each exposure was conducted on the same piece of simulated road for consistency. Data for 22 participants was captured for the familiarisation sessions.

The following results were identified from the familiarisation sessions.

5.1.1 Familiarisation 1 – 40 mph:

At 40mph, most participants estimated the speed of the vehicle to be lower than the actual speed with an average predicted speed of 31mph. Only one participant (5%) rated it higher (50mph) and three participants (14%) scored it correctly at (40mph). The remaining 18 (81%) scored it lower, with 4 participants (18%) scoring it considerably lower at 20mph.

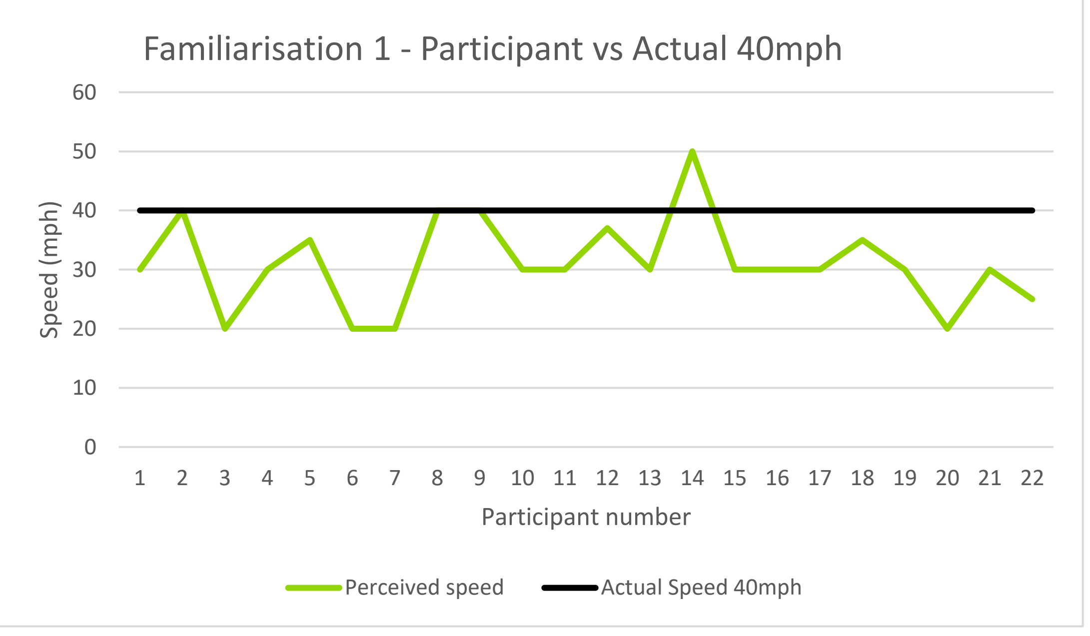


Figure 16 - Spread of participant estimated speed vs. actual at 40mph

5.1.2 Familiarisation 2 - 50 mph:

As for the 40mph scenario, the majority of participants believed that they were travelling slower than the simulated 50mph, with an average of 38mph. Again, only one participant (5%) rated the experience as faster (a rating of 60mph) and only 2 participants (9%) rated the experience the same as the simulated speed of 50mph. The remaining 19 participants (86%) of participants rated the experienced slower, with 7 participants (32%) reporting that the experience felt like 30mph, 20mph lower than the simulated speed.

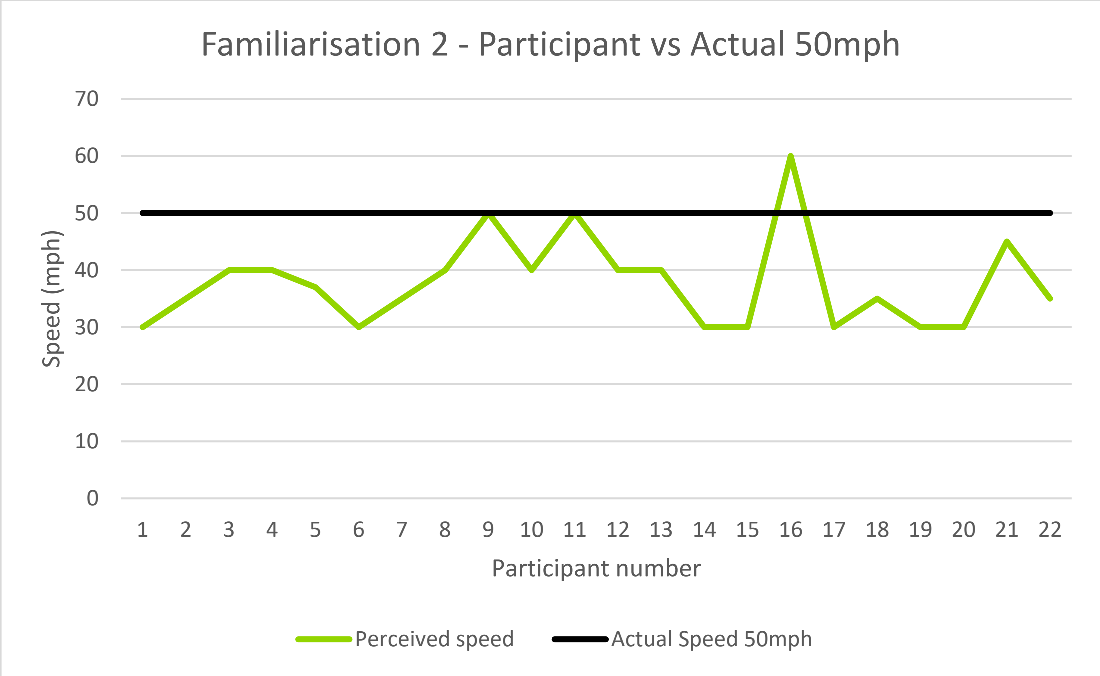


Figure 17 - Spread of participant estimated speed vs. actual speed at 50mph

5.1.3 Familiarisation 3 - 60 mph:

The fastest simulated speed had the largest number of incorrect scores, with only one person (5%) accurately scoring the speed and one participant (5%) judging the speed as faster at (65mph). The remaining 20 participants (91%) rated the speed as lower, with an average speed of 41mph across all participants, which is 19mph lower than the simulation. One participant (5%) rated the speed at 25mph, which is 35mph lower than the simulated experience, with five participants (23%) rating it at 30mph (30mph lower).

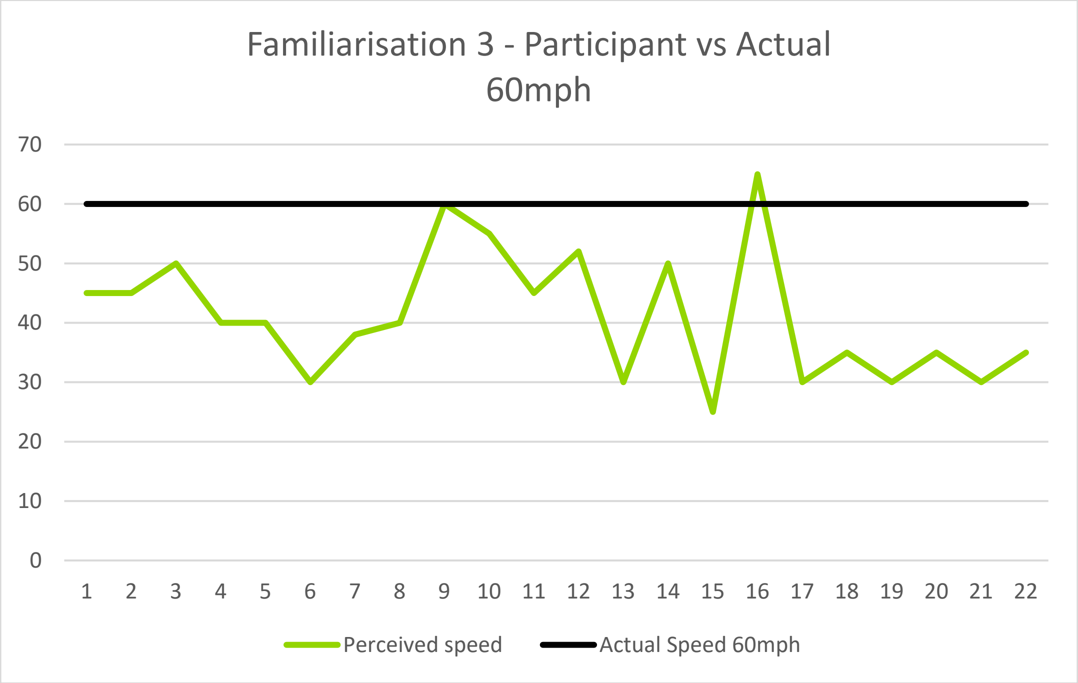


Figure 18 - Spread of participant estimated speed vs actual speed at 60mph

In summary, it can be concluded that participants, on average, rated the experience slower than the simulated speed in all conditions. The faster the experience, the greater the level of disparity in scores and a generally greater level of error. The results highlight that the overall experience that the participants had, led to an overall lower perception of the sensation of speed and as a result, their perceived comfort and feedback from the scores should take this into account.

6 Comfort Results

This section presents a summary of the comfort data. To see the full data and charts created for the analysis process, please see Appendix A. To see the participant demographics analysis, please see Appendix B.

6.1 Overall Scenario Rating Summary:

The following table is a summary of the overall comfort ratings given by participants and the overall rating of the comfort Algorithm.

The table shows the most common comfort rating assigned by the 27 participants. The comfort model shows the Overall Comfort rating, or a FAIL with the projected rating in ( ). The lowest comfort factor score produced by the comfort algorithm is shown in the grey column.

Scenario	Participant Rating	Comfort model Rating	Comfort Model's lowest factor rating
1	B	FAIL (A)	B
2	A	A*	A*
3	B	FAIL (A)	D
4	A-C	FAIL (A*)	B
5	A	A*	A*
6	A	FAIL (A*)	A
7	C	A*	A*
8	A	FAIL (A*)	D

Figure 19: Shows the overall comfort ratings of participants, the comfort algorithm and the lowest comfort factor rating generated by the algorithm.

6.2 SCENARIO COMFORT DATA ANALYSIS

The following section breaks down the data / results per scenario.

6.3 Scenario 1: Motorway merge

To watch a video of the Scenario and the comfort model click here: LINK

There was divergence in the overall comfort rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model' lowest factor rating
1	B	FAIL (A)	B

The reasons for a failure result by the comfort algorithm was due to the following:

- Aggressive acceleration at the beginning of the scenario
- Headway too small to the car ahead in the slow lane.

6.3.1 Participant Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
2	Extremely comfortable	7%	A*
8	Very comfortable	30%	A
10	Comfortable	37%	B
5	Acceptable	19%	C
2	Uncomfortable	7%	D
0	Very Uncomfortable	0%	F
27			

Figure 20: Shows the participant overall comfort rating for Scenario 1

6.3.2 Participant individual comfort factors:

Negative comfort factor ratings (Aggressive & Very Aggressive), by participants were:

Headway (3) 12%,

- Vehicle proximity (1) 4%,
- Pulling out of the junction (2) 8%
- Lane positioning (1) 4%.

6.3.3 Verbatim analysis:

This scenario involved merging onto a motorway, and the most common theme from qualitative comments were that there were no wing mirrors and as a result participants couldn't see behind them. Some participants voiced that this felt limiting and they tried to move their head more to check the inside lane, as they ‘couldn’t see cars coming/overtaking till they were next to you’. This lack of situational awareness was cited as the reason that some participants felt uncomfortable, and there was a comment made that ‘it used indicators, but wanted to verify myself’ and ‘having a mirror or screen would help comfort’. There were some comments stating that the ego vehicle ‘got close to the car ahead’ and some participants reported pulling the trigger as a result.

6.4 Scenario 2: Motorway lane change and Exit

To watch a video of the Scenario and the comfort model click here: **LINK**

There was good alignment in the overall rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model's lowest factor rating
2	A	A*	A*

The comfort algorithm aligned well with the participant's ratings with 70% of participant's rating the scenario as Comfortable, Very comfortable or Extremely comfortable.

6.4.1 Participants Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
3	Extremely comfortable	11%	A*
10	Very comfortable	37%	A
6	Comfortable	22%	B
5	Acceptable	19%	C
3	Uncomfortable	11%	D
0	Very Uncomfortable	0%	F
27			

Figure 21: Shows the participant overall comfort rating for Scenario 1

6.4.2 Participant individual comfort factors:

Negative comfort ratings (Aggressive & Very Aggressive), by participants were:

- Deceleration (4) 15%,
- Vehicle proximity (3) 12%,
- Pulling out of the junction (1) 4%,
- Changing Lane (8) 31%.

6.4.3 Verbatim analysis:

This was changing lanes on a motorway and then taking a motorway exit. There was a mixed spread of themes with no clear prevalent theme. Participants largely summarised what occurred during the scenario, commenting that there was 'no change in acceleration or deceleration as it pulled off' and that they 'felt comfortable at all times, didn't pull the trigger'.

A small number of participants commented on the following:

- the lane change occurred too soon after the car in the inside lane was passed
- The lane change manoeuvre could have been smoother
- The lack of mirrors on the Ego did not enable participants to check that the ego had cleared the car in the slow lane before moving into the slow lane
- The slip lane on the motorway was not long enough with no signage to show the exit was approaching
- The width of the slip lane was not appropriate for the speed.

6.5 Scenario 3: Roundabout no traffic

To watch a video of the Scenario and the comfort model click here: **LINK**

There was divergence in the overall comfort rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model's lowest factor rating
3	B	FAIL	D

The reason for the comfort algorithm failure in the scenario was due to the following:

1. Excessive speed on the curve leading up to the roundabout (Lateral Accel)
2. Excessive when turning on to the roundabout and the speed on the roundabout (Lateral Accel)

6.5.1 Participants Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
0	Extremely comfortable	0%	A*
2	Very comfortable	8%	A
9	Comfortable	35%	B
6	Acceptable	23%	C
8	Uncomfortable	31%	D
1	Very Uncomfortable	4%	F
26			

Figure 22:Shows the participant overall comfort rating for Scenario 3

6.5.2 Participant individual comfort factors:

Negative comfort factor ratings (Aggressive & Very Aggressive), by participants were:

- Acceleration (5) 19%,
- Deceleration (8) 31%,
- Pull out of junction (20) 77%,
- Lane positioning (6) 23%,
- Changing Lane (3) 12%

6.5.3 Verbatim analysis:

Here the AV joined a roundabout and then exited on the second exit with no other cars present. Participants' most prevalent comments were that they were uncomfortable with the joining of the roundabout (85% of participants failed to identify the scenario as a roundabout, mistaking it for a curved road with 2 junctions). They perceived the turns as 'very sharp', 'very robotic and spinning on the spot'. Some participants also expressed surprise that it didn't come to a stop at the junction and that they felt like it 'went almost on the other side of the road'.

6.6 Scenario 4: Roundabout with traffic

To watch a video of the Scenario and the comfort model click here: LINK

There was divergence in the overall comfort rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model's lowest factor rating
4	A-C	FAIL	B

The reason for the comfort algorithm failure in the scenario was due to the following:

1. Insufficient gap to the car on the roundabout

6.6.1 Participants Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
1	Extremely comfortable	4%	A*
8	Very comfortable	31%	A
7	Comfortable	27%	B
8	Acceptable	31%	C
2	Uncomfortable	8%	D
0	Very Uncomfortable	0%	F
26			

Figure 23:Shows the participant overall comfort rating for Scenario 4

6.6.2 Participant individual comfort factors:

Negative comfort factor ratings (Aggressive & Very Aggressive), by participants were:

- Deceleration (2) 8%,
- Vehicle Proximity (4) 15%,
- Pull out of junction (11) 42%,
- Lane positioning (1) 4%,

6.6.3 Verbatim analysis:

As with scenario 3 this involved joining a roundabout and taking the 2nd exit. There were also 2 cars on the roundabout. Participants stated they were happy with the speed that the ego vehicle travelled at, but they expected it to stop and give way to the oncoming vehicle and it didn't. They stated that 'I would have waited for the car to pass before I would have pulled out', and 'not sure I would have gone ahead'. Some recognised that it was a very similar scenario to Scenario 3, but they felt the turning was 'smoother' and that 'it didn't come out as far into the road as it went round the corner' (as the order was counterbalanced, this would only apply to those who experienced Scenario 3 before Scenario 4).

6.7 Scenario 5: A Road oncoming car in adjacent lane

To watch a video of the Scenario and the comfort model click here: **LINK**

There was good alignment in the overall comfort rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model's lowest factor rating
5	A	A*	A*

6.7.1 Participants Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
4	Extremely comfortable	15%	A*
10	Very comfortable	38%	A
9	Comfortable	35%	B
2	Acceptable	8%	C
1	Uncomfortable	4%	D
0	Very Uncomfortable	0%	F
26			

Figure 24: Shows the participant overall comfort rating for Scenario 5

6.7.2 Participant individual comfort factors:

Negative comfort factor ratings (Aggressive & Very Aggressive), by participants were:

- Vehicle Proximity (4) 15%,
- Lane positioning (3) 12%

6.7.3 Verbatim analysis:

This scenario involved driving on an A road and passing an oncoming car, after which participants summarised that they were happy with the speed of travel. Participants observed that there was a car coming the other way but stated that overall, the scenario was ‘uneventful’ and a ‘nice comfortable easy drive’. A small number of participants found the second curve in the road as ‘sharp’ and one pulled the trigger as a result of this.

6.8 Scenario 6: A Road, slowing down for slow car ahead.

To watch a video of the Scenario and the comfort model click here: **LINK**

There was divergence in the overall comfort rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model's lowest factor rating
6	A	FAIL	A

The reason for the comfort algorithm failure in the scenario was due to the following:

1. Excessive acceleration and Jerk at the beginning of the scenario as the ego accelerated to the target speed.

6.8.1 Participants Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
4	Extremely comfortable	15%	A*
9	Very comfortable	35%	A
7	Comfortable	27%	B
4	Acceptable	15%	C
2	Uncomfortable	8%	D
0	Very Uncomfortable	0%	F
26			

Figure 25:Shows the participant overall comfort rating for Scenario 6

6.8.2 Participant individual comfort factors:

Negative comfort factor ratings (Aggressive & Very Aggressive), by participants were:

- Acceleration (1) 4%
- Deceleration (5) 19%
- Pull out of junction (2) 8%
- Lane positioning (2) 8%.

6.8.3 Verbatim analysis:

Here the ego vehicle was driving on an A road, following a car ahead and then slowed down due to a slower car ahead. Participants stated that they felt there was enough distance to the car in front, though many felt that the approach to it was a bit fast and one pressed the trigger as they felt 'there was a bit of a delay before my car started to slow down' and 'approached car in front too quickly'.

6.9 Scenario 7: T Junction no traffic

To watch a video of the Scenario and the comfort model click here: **LINK**

There was good alignment in the overall rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model's Lowest factor rating
7	C	A*	A*

The comfort algorithm's overall rating and that the participant's ratings were not well aligned.

The comfort factors that received negative comfort ratings were:

- Deceleration: Aggressive: 11,
- Pulling out of junction: Aggressive: 12, Extremely Aggressive: 6

6.9.1 Participants Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
1	Extremely comfortable	4%	A*
4	Very comfortable	15%	A
6	Comfortable	23%	B
11	Acceptable	42%	C
4	Uncomfortable	15%	D
0	Very Uncomfortable	0%	F
26			

Figure 26: Shows the participant overall comfort rating for Scenario 7

6.9.2 Participant individual comfort factors:

Negative comfort factor ratings (Aggressive & Very Aggressive), by participants were:

- Acceleration (4) 15%,
- Deceleration (11) 42%,
- Pull out of junction (18) 69%,
- Lane positioning (3) 12%,
- Changing lane (2) 8%,

6.9.3 Verbatim analysis showed

This involved turning right at a T junction with no other cars, however the most frequent feedback from participants was that it was an uncomfortable turn that felt 'not very natural' and 'not human like'. Some participants reported having pulled the trigger. Some stated that they expected the car to stop and that it pulled off while they were checking for oncoming cars at the junction.

6.10 Scenario 8: T Junction with traffic

To watch a video of the Scenario and the comfort model click here: **LINK**

There was good alignment in the overall rating:

Scenario	Participant Rating	Comfort model Rating	Comfort Model's lowest factor rating
8	A	FAIL	D

The reason for the comfort algorithm failure in the scenario was due to the following:

1. Excessive acceleration and Jerk at the beginning of the scenario as the ego accelerated and then decelerated towards the T Junction.

6.10.1 Participants Overall comfort ratings:

Responses No.	Overall Comfort Rating	Percentage	Comfort Rating
2	Extremely comfortable	8%	A*
10	Very comfortable	38%	A
9	Comfortable	35%	B
1	Acceptable	4%	C
4	Uncomfortable	15%	D
0	Very Uncomfortable	0%	F
26			

Figure 27: Shows the participant overall comfort rating for Scenario 8

6.10.2 Participant individual comfort factors:

Negative comfort factor ratings (Aggressive + Very Aggressive), by participants were:

- Acceleration (1) 4%,
- Deceleration (4) 15%,
- Headway (3) 12%,
- Pull out of junction (4) 15%,

6.10.3 Verbatim analysis showed

Here the vehicle turned right at a T junction with cars approaching the T junction in both directions, and participants overwhelmingly reported that they expected it to stop and it didn't. The deceleration was questioned with comments that it was 'aggressive' and 'hit the brakes quite early'. Some participants stated that they didn't pull the trigger as they 'felt fine'.

Participants commented that compared to Scenario 7, the turning was smoother. A few participants felt uncomfortable when merging into the gap between the cars coming from the left. A few participants commented on the headway at the end of the scenario because the ego was catching the car ahead.

7 Discussion

7.1 Comfort Factors

7.1.1 Acceleration & Jerk

Analysis of the results indicates that one of the primary reasons for the lack of alignment between the comfort algorithm and participant's comfort ratings, was the lack of motion/acceleration simulation. This was one of the known drawbacks of switching from using a motion platform/driving simulator with a wide angled field of view VR headset, to a mobile VR headset strategy and delivering the trial remotely in participant's homes.

The activity that provides evidence to this assumption is the familiarisation activity conducted at the beginning of the user trial. Participants were shown a scenario where the Ego was traveling down an A road at 3 set speeds (40, 50 and 60mph). Participants were asked to estimate how fast they thought the Ego was travelling. The results showed significant divergence between the Ego' speed and the participants perceived speed.

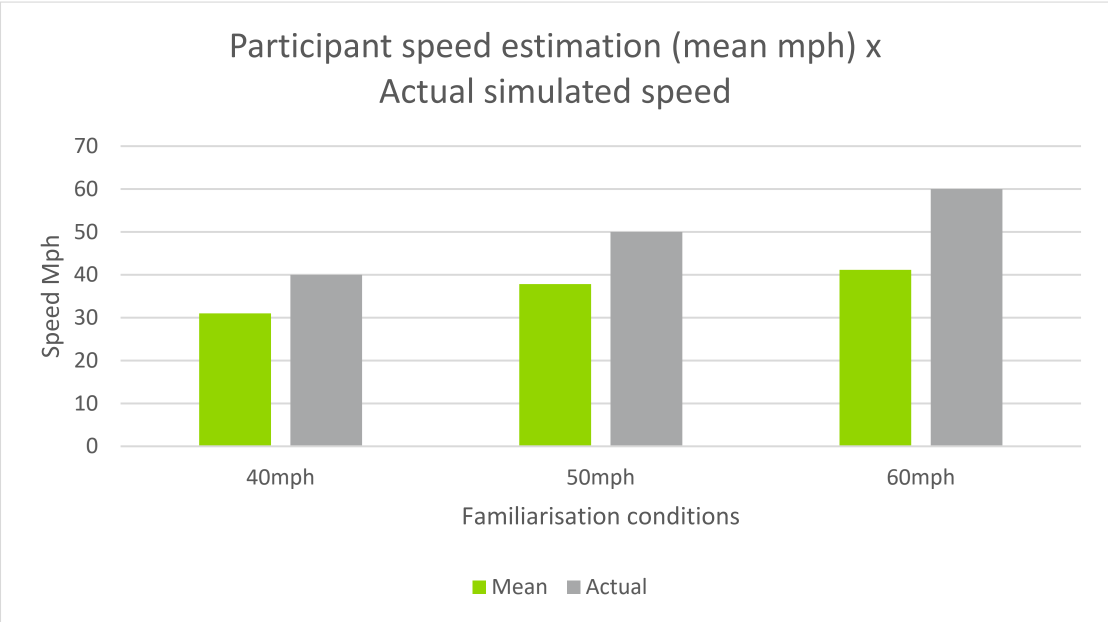


Figure 28: Shows the divergence between the actual speed and the speed perceived by participants

The familiarisation/speed calibration activity appears to confirm the literature that concludes that peripheral vision has a significant influence in correctly identifying speed and acceleration/deceleration.

Analysis of the user trial's verbatim showed that participants noticed deceleration most often, but not exclusively, when the Ego was approaching a junction or when there was a vehicle directly ahead. This may have been due to them fixating/concentrating on a potential hazard ahead that was in their direct line of sight.

The other consideration is the way the scenarios were generated from the suite of software used by the VeriCAV project. The user trial scenarios were short in duration (30-70seconds) and sometimes it was necessary for the scenario to begin with the Ego stationary.

This led to the participant experiencing an immediate acceleration at the beginning of the scenario. It is not known how much acclimatisation is necessary to be able to detect acceleration solely through the visual senses and if situational awareness contributes, but it is more likely that peripheral vision and motion simulation has a greater affect.

Furthermore, it is well known that if objects in the environment and are in close proximity, such as trees, signs and buildings, this has an impact on the perception of speed.

The alternative consideration that may have caused a comfort rating miss alignment, was an overly punitive scoring system (the negative scoring applied for each comfort factor boundary infringement). For each second the data exceeded the comfortable thresholds, the negative scoring was:

- -1/sec (normal),
- -10/sec (aggressive)
- -20/sec (very aggressive)

7.1.2 Headway

The intention of the user trials was to evaluate headway in Scenario 6 (an A Road scenario with car slowing/slow ahead). However, headway was also a factor in Scenario 1 (Motorway merge). The researchers did not realise that the headway thresholds had been exceeded until after the final simulations were completed. Scenario 1 was therefore and unplanned, but useful additional headway condition to assess.

The comfort model recognised the headway to the car ahead, towards the end of Scenario 1, was less than the 'Comfortable' threshold and therefore applied negative scoring. The comfort model did not record any headway comfort threshold infringements in Scenarios 6, however there were negative comments identified in the verbatim

- Participants indicated that a more progressive speed reduction would have been preferable, especially due to car ahead being clearly visible.

It should be noted that Apollo was only used for Scenarios 4 and 8. For the other scenarios, OpenScenario was used to co-ordinate the movements of the Ego and other vehicles.

OpenScenario does not define gaps/distances to other vehicles unless it is programmed to do so and any accelerating/braking behaviour is simplistic. This would explain why the Ego got too close to the car ahead in Scenario1 and why the braking due to a slower car ahead in Scenario 6 was abrupt. That said, the results from the comfort model and those of participants should be aligned, but that was not the case.

In Scenario 1, 3 participants rated the headway as aggressive and 4 participants verbatim noted the proximity of the car ahead. The reducing headway did not negatively affect the overall comfort rating. The comfort model failed the scenario due in part to the headway separation/gap.

The reason that participants did not fail the scenario due to a reduced headway, may be due to the reduced headway occurring at the very end of the scenario. Some participant's verbatim indicate that because the video faded out as the Ego approached the vehicle in the lane ahead, they could not determine/conclude if the Ego would continue to approach the car ahead, slow down or overtake it.

7.1.3 Vehicle Proximity and Lane positioning

In lane positioning of the Ego and the width of lanes were such that there was no infringement of the comfortable comfort rating of >0.8m. Therefore, the researchers did not expect to see participants negatively rating the Ego's lane positioning.

However, there were a few participants who rated the Ego's lane positioning as Aggressive or Very Aggressive. Looking at the participants demographics showed that these participants had low sensation seeking scores and a high locus of control. The literature shows that speeds, lane width, environmental conditions and vehicle size have an effect on vehicle in the lane. In real world driving, drivers perform small adjustments to their lane positioning when confronted with oncoming vehicles. This is often to satisfy a personal desire for a larger safety margin or to make a trajectory that is more comfortable. The simulation software does not account for this as the lanes were defined widths.

This user trial did not have the time to explore a greater number of road widths, vehicle to vehicle proximity (adjacent) or vehicle types, but the simulation software built for this user trial could be easily adjusted/enhanced to enable this research and further test the comfort model's lateral offset comfort boundaries.

7.1.4 Pulling out of junction (Gap acceptance)

Two of the eight Scenarios tested gap acceptance. These were the Scenario 4 (T junction with traffic) and Scenario 8 (Roundabout with traffic).

Scenario 8 (T Junction with traffic) used a gap of 10 seconds between the vehicles coming from the left. The majority of participants were happy which aligns to the literature. Three participants rated the pulling out of the junction as 'Aggressive' and one rated it as 'Very Aggressive'. Allowing for normal distribution it is expect that drivers with lower risk thresholds will hesitate/want a gap greater than 10 second. Also, taking into consideration that there was no traffic after the 2 cars from the left, there is a natural instinct to wait for a larger gap if the time needed to wait at the T Junction is reasonable.

In Scenario 4 (Roundabout with traffic) participants did not rate the behaviour of the Ego as aggressive, even though the Gap from their vehicle to the oncoming car on the roundabout was only 3 seconds. The model predicted that a gap of 3 seconds should result in an Extremely Aggressive comfort rating.

When considering that the data used to inform the comfort model was taken from T Junction data analysis, it may be reasonable to conclude that the data was not appropriate to use for roundabouts. The speeds of vehicles on roundabouts tend to be slower than the traffic traveling on a carriageway where there are T Junctions. Furthermore, when drivers approach a roundabout they tend to slow down gradually and creep towards the entry point of the roundabout. This allows the driver to accelerate to a speed to match those vehicles on the roundabout faster and therefore negotiate and merge into smaller gaps.

7.1.5 Changing lanes

Scenario 1 and 2 included lane change manoeuvres. The comfort algorithm measured lane change parameters by looking at the combined Acceleration and Jerk data.

Participants could not experience Acceleration and Jerk due to the aforementioned lack of a motion platform to simulate vehicle motion/dynamics. Instead participants rated the Ego's lane change manoeuvres through observation of the 360 video in the VR headset. As a result, participants were solely using visual observation to rate the vehicle's positioning, speed, timing and proximity to other vehicles.

Analysis of the verbatim showed that some of the participants, wanted to check their mirrors and blind spot when joining the motorway from the slip road (Scenario 1), and when moving into the slow lane having overtaken a vehicle in the slow lane (Scenario 2). Because they could not check the proximity and positioning of vehicles, they had to trust the system to perform the manoeuvre. The comfort ratings were generally positive (Comfortable or Normal) with the observation that they would have liked to check in their mirrors or look over their shoulder towards their blind spot (neither of which were possible with the simulation).

In Scenario 2, participants had context/knowledge of the position of the other vehicles in the lanes ahead before overtaking and then pulling into the slow lane. As a result, 8 participants rated the lane change manoeuvre as Aggressive or Very Aggressive. The verbatim analysis of these participants showed that they felt that the Ego had pulled over too soon/close the overtaken vehicle. Some participants commented that there was no need to pull over so soon and aggressively because there was plenty of space and time to pull into the slow lane over a greater distance/time.

These comments are interesting from a point of view of headway and link to the roundabout verbatim. Drivers are generally concerned about their proximity to other vehicles in all directions and tend to maximise the gap to other vehicles where possible. It therefore could be useful for ADS engineers to consider a bubble that surrounds the vehicle in all directions. This bubble may be larger ahead of the vehicle for braking/gap purposes, but may also be larger than necessary (from an engineering point of view), rearwards and sidewards to account for the comfort of other road users.

The verbatim and comfort ratings raise an interesting question about what information passengers desire when travelling in a CAV, where it is presented and how often this information would be consulted by passengers. It is reasonable to speculate that first time users would be interested, especially if they are experienced drivers, to know where the other road users are and have acknowledgement that the CAV has recognised their presence. It is also reasonable to speculate that this desire for vehicle proximity information will reduce over time/usage as passengers gain confidence in the capabilities and behaviours of CAVs.

7.2 Comfort Model review

The table shows the most common comfort rating assigned by the 27 participants. The comfort model shows the Overall Comfort rating, or a FAIL with the projected rating in ( ). The lowest comfort factor score produced by the comfort algorithm is shown in the grey column.

Scenario	Participant Overall Comfort Rating	Comfort Algorithm Overall Rating	Comfort Algorithm's lowest Comfort Factor Rating
1	B	FAIL (A)	B
2	A	A*	A*
3	B	FAIL (A)	D
4	A-C	FAIL (A*)	B
5	A	A*	A*
6	A	FAIL (A*)	A
7	C	A*	A*
8	A	FAIL (A*)	D

Figure 29: Shows the overall comfort ratings of participants, the comfort algorithm and the lowest comfort factor rating generated by the algorithm.

Overall, the comfort Algorithm performed well. Based on its criteria, its scoring system and the data it analysed, the results and the scores awarded were deemed to be appropriate. However, the overall ratings, whether the ADS Passed or failed, were A or A* which does not provide much granularity.

Based on the participant's overall comfort ratings and verbatim analysis, when an uncomfortable Ego behaviour was experienced, this affected the respective comfort factor and the overall comfort score. Because acceleration/Jerk was not simulated/experience by the participants, it is not possible to determine the effect on their scores. However, it is reasonable to expect their comfort ratings to go down if they physically experience acceleration/jerk.

Instead of averaging the sum of the comfort factor's score/rating, it is proposed that the overall comfort rating should duplicate/mirror the lowest comfort factor score/rating.

This approach makes sense due to the following:

- When an ADS controls the Ego in an uncomfortable manner, this could be for a short/momentary period of time, but the uncomfortable behaviour has an effect on the overall perception of comfort (passengers)
- If more comfort factors are added to the Algorithm, the current system will mask the negative/uncomfortable behaviours if the average of the individual comfort factor's scores are used
- If scenario lengths were to be extended (>70seconds), this would further mask momentary negative/uncomfortable ADS behaviour.

It is not possible to conclude definitively that the punitive/negative scoring system for the comfort factors was appropriate. In part this is due to experimental method's inability to replicate Acceleration and Jerk (no motion platform). Acceleration and Jerk are considered, by the researchers in this project, to be the most influential factors contributing to comfort. Therefore, because participants were not able to experience the dynamics of the vehicle, it is impossible to determine how their comfort ratings would be affected.

As part of the data analysis, the graphs produced to visualise the comfort algorithm's comfort thresholds and the negative scoring system helped the researchers to determine the rate of acceleration. For example, at the beginning of Scenario 1, the Ego accelerated from 0-60 in 6 seconds. This was a rapid acceleration that most passengers would find uncomfortable in a straight line. This acceleration was combined with a curve in the road and therefore it can be concluded that this experience would have been very uncomfortable. It is therefore reasonable to determine that the scoring system which resulted in 0 points in that 10 second segment was appropriate.

It will require further evaluation to determine the validity of a comfort factors scoring <60points /10 second period to result in an overall test Fail. To validate the Fail criteria, it is proposed that a scenario is generated, within which a range of exposure times to a fixed level of discomfort are varied. Participants could then be asked to rate the experience and validate if the negative scoring/comfort ratings are appropriate.

7.3 Participants

Participants were recruited from within the workforce of the Connected Places Catapult. Due to health and safety considerations and the value of the VR equipment being used, it was not appropriate to recruit members of the public. Despite the restrictions imposed on the pool of potential participants, the demographics of the participants was balanced in gender, age and driving experience.

It was surprising to the researchers how well the participants, despite their age and technological competencies, were able to use the VR equipment. In part this was helped by the creation of short tutorials for each pertinent aspect of the VR headset. This helped the participants, many of whom had never used a VR headset, to go through each step of the adjustment of the fitting process, hand controller functionality and space setup/configuration in their homes. This resulted in the trial facilitators only needing to offer minor guidance and support during the trial, helping maximise the time available for the trial and minimise distraction from the research objectives.

7.4 Experimental method

Considering the constraints of running a user trial during the 3rd Covid lockdown, the experimental method was successful in collecting the necessary data. In fact, the user trial and bespoke software tools created by Connected Places Catapult, collected more data than the team was able to analyse in the time remaining in the overall VeriCAV project. This data could be of interest to academic institutions and should be explored after the project end date, because there are more insights that can be drawn out of the data. Furthermore, the data analysis tools could be of use within Connected Places Catapult or by external institutions to support new experimental research and analytical methods.

The online survey tool (Survey Monkey) proved to be highly effective tool throughout the planning, delivery and data analysis processes. Going fully digital/online minimised the time needed to transcribe handwritten notes and provided researchers the ability to extract data in different formats to support our specific data analysis requirements.

Microsoft's Teams video conferencing software was used to facilitate the trails and proved to be an effective communication tool. The tutorial videos and written instructions were very effective, without which running the trial using would have been a lot more challenging. The primary issue with facilitating a VR experiment is not being able to see what the person wearing the VR is seeing. This is challenging when the person wearing the headset is physically next to you, let alone when they are not physically present.

The participants needed to have a common terminology and a basic understanding of the controls interfaces to be able to relay/explain any issues they are experiencing to the trial facilitator. Without this baseline knowledge and experience, the trials could have been a lot more challenging to deliver.

The questionnaires used in the data collection process were effective and delivered the data the researchers needed. However, in hindsight, it would have been useful to customise the in-trial questions asked, based on the comfort factors being targeted in each scenario. There was also ambiguity with the question "How comfortable were you when your vehicle changed lane". Some participants determined that changing lane could apply to Scenario 1 when the Ego joined the motorway from the slip lane. Participants were also unsure about its applicability to the T-junction Scenarios (7&8).

The question "How comfortable were you when you pulled out? (Roundabout/T-junction) was also slightly ambiguous. The act of pulling out has multiple components that could contribute to comfort, such as the trajectory, the rate of acceleration, other vehicles present etc. It would be useful to have subset questions to the main questions, that would appear if a negative comfort response were given. This would enable the delivery of the questionnaires to be streamlined if the comfort levels are high and collect details about contributary factors if the comfort rating levels are low.

7.5 Ego Vehicle / Other Vehicle Behaviour

The project's software development and engineering teams had many challenges to overcome in the overall VeriCAV project, part of which was the integration of software tools from different vendors. This integration process required significant development time and refinement. This had an effect on the time available to implement the software developed for the user trials. This affected many of the critical components necessary to deliver the simulations required for the user trial. In addition to these challenges, halfway through the development process, the user trials were moved from office based to remote/home based.

These software development challenges meant that the software took longer to develop and did not meet all the requirements.

As a result, it was only possible to use the ADS (Apollo) in Scenarios 4 and 8 (Roundabout and T-Junction with traffic). The Ego and the other vehicles trajectories for the other scenarios were determined by the OpenScenario files..

The difference between the Apollo ADS controlling the Ego and OpenDrive is as follows:

- Apollo = non determinative simulation

    - The starting point and end points are defined. The ADS works out how to get to the destination.
- OpenDrive =determinative simulation
    - The starting point, the trajectory, speeds and end point are determined. The Ego follows the instructions and can not deviate.

Using OpenDrive and having trajectories defined by splines, results in movements that are not 100% natural when experienced from the seat of the virtual Ego vehicle. A good way to think about this is to think about a Scalextric toy. There are grooves in the track that the toy car has to follow. As a result, any imperfections in the trajectory/spline are translated into movements in the car that can be detected by the participant.

There were noticeable effects on the comfort ratings, most notably in Scenario 3: Roundabout no traffic. This was primarily due to tight trajectories at the entry point of the roundabout. This caused the car to rotate on the spot instead of following a natural turning trajectory, where the rear wheels follow the front wheels at an offset angle.

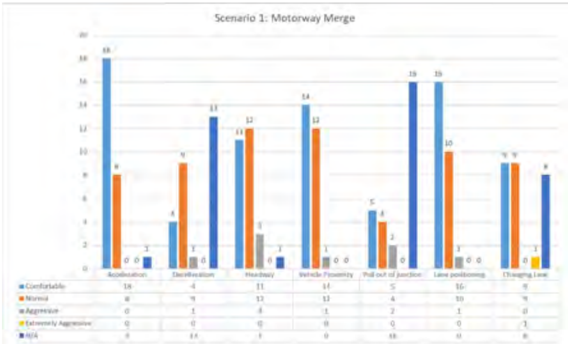
Despite the known limitations, it was possible to create realistic behaviours and the feedback from participants on the simulation's realism was positive. A further positive aspect of using a mixture of OpenDrive and Apollo was that there were 2 extremes for participants to experience. The participants favoured the Apollo scenarios, resulting in generally higher comfort factor ratings.

Another aspect that needs consideration is how the roads are built in OpenDrive. It is important to use road geometries that are representative and conform to regulations. For example, the roundabout scenarios (Scenario 3 and Scenario 4) had entry/exit roads that were not appropriate due of the angle the lanes relative to the diameter of the roundabout. As a result, there was very tight angle to negotiate if the Ego was to stay in lane. This affected the speed and turning behaviour of the Ego as it entered and exited the roundabout.

7.6 Statistics and data visualisation

The data visualisations produced were useful for the user trials data analysis. The graphs, tables and analysis tools enabled researcher used to analyse ADS/Ego behaviour from a macro and micro level. See Figure 30: Shows a selection of the data analysis and visualisation tools created for data analysis.Figure 30 below.

For example, using the comfort algorithm summary chart (top right), it is possible to identify in which 10 second segment a failure occurred. With the scenario review tool (bottom right), which contains the dynamic graphs and an in-cockpit video replay, it is possible to review the vehicle's dynamics, the comfort model thresholds and any negative scoring incurred. This can be reviewed second by second alongside the scenario video to identify which comfort factors exceeded the comfort thresholds and observe the negative score accumulation.



Comfort Algorithm Scores: Scenario 1					
	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars
Seconds	CF1A	CF1B	CF2	CF3	CF4 : Not measured!
10	0	15.05	100	100	100
20	99.55	99.75	100	100	100
30	100	100	100	100	100
40	100	100	100	90.8	100
50	100	100	100	18.7	100
Averages	79.91	82.96	100	81.9	100
Comfort Grade	B	A	A*	A	N/A
Pass/Fail	FAIL	FAIL	PASS	FAIL	N/A
Overall comfort score	86.1925				
Overall Comfort PASS/FAIL	FAIL	A			

Participant Results				Algorithm Results			Scoring Matrix		
No. of	Percentage	Overall Comfort Rating		Comfort Score	Pass/ Fail		Algorithm Grades	Algorithm Ratings	Overall Ave Score
2	8%	Extremely comfortable					A*	Extremely comfortable	>90
10	38%	**Very comfortable**					A	Very comfortable	>80
9	35%	Comfortable					B	Comfortable	>75
1	4%	Acceptable					C	Acceptable	>65
4	15%	Uncomfortable					D	Poor	>60
0	0%	Very Uncomfortable			FAIL		F	Fail / uncomfortable	<60
26									

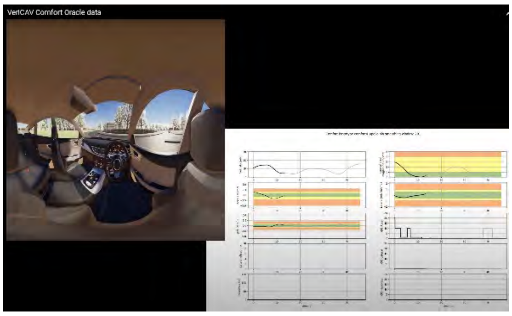


Figure 30: Shows a selection of the data analysis and visualisation tools created for data analysis.

With some further refinement, these tools should be able to support engineers to identify and rectify ADS behaviours that do not align to the comfortable and normal comfort levels.

When considering the future potential applications of the comfort model, it is worth considering how the comfort algorithm and the above data visualisation could be integrated into the VeriCAV comfort Oracle. A data dashboard for each test algorithm would be a logical, practical and useful feature to develop and would serve to inform CAV / ADS developers about how to improve functionality. Conducting more development in simulated environments prior to real world testing will prove to be cost effective and should reduce development time in the physical world.

8 Conclusion

Data analysis showed that the VeriCAV Comfort Algorithm has strong potential as a method for defining ADS comfort in a simulation. However, more development and testing is required to verify the Comfort Algorithm. Ideally this would involve re-running the user trial using a motion platform to replicate Acceleration and Jerk and using a larger number of participant recruited from the general public. It would also be highly beneficial to work with an OEM/ADS development company to review the algorithm, software tools, processes and data visualisation to create a product that meet the needs of CAV/ADS engineers.

8.1 Comfort Algorithm vs. Participant Comfort Ratings

There was partial alignment between the comfort algorithm and participant Comfort Factor ratings. There are a number of factors to consider:

- Comfort is highly subjective, personal and contextual. The researchers did not expect direct and consistent alignment, but instead were looking for trends/evidence in the data
- It was not possible to simulate motion (Acceleration/Jerk). This created divergence between the participant and the algorithm's comfort ratings:
    - Participants were not able to detect physically, or observe (via the VR headset)
    - The comfort Algorithm was able to measure and quantify
- Due to a small sample size (26), a limited distribution of participants and a large number of variables, it was challenging to make definitive conclusions within the budget, time and resource available
- Using VR to expose participants to simulated scenarios was expected to have an influence on the results especially as many of the participants had never used VR before
- The participants demographics shows an equal gender split, a wide spread of ages, years of driving experience, attitudes to technology, Automation and CAVs
    - To manage any negative effects of using VR technology, researchers created tutorial videos and gave participants time to familiarise themselves with the hardware prior to the user trial. All participants rated the usability of the system as "Easy" or "Very Easy".

Excluding Acceleration and Jerk, participant's comfort ratings for Headway, lateral offset and Gap acceptance (involving visual perception/ judgment), were much closer aligned to the scores/ratings of the Comfort Algorithm.

8.2 Overall Comfort Rating

The algorithm's overall comfort rating did not function as intended, with overall Comfort Ratings scoring either an A* or a Fail (with a potential of an A/A*).

Data analysis indicates that a more appropriate method for rating the overall comfort of an ADS would be for the comfort algorithm to duplicate the lowest individual comfort factor rating.

This aligns to the notion that a negative comfort experience/instance within a scenario will affect the overall comfort rating. As comfort is personal, perceived on a continuum and constantly changes along a scenario, no two journeys will necessarily be rated the same. Further research is necessary to determine if the algorithm's calculations/scoring methodology would still be appropriate if the comfort algorithm evolves to include more comfort factors or if the scenario duration is increased beyond the 70 seconds maximum scenario exposure time participants experienced in this user trial.

8.3 Summary

Throughout this project, researchers have been challenged to define comfort, both theoretically and objectively. The scientific literature highlights this challenge and there are many competing models on the topic, with no clear path to a unified concept/definition of comfort.

When analysing the data and speaking to the participants, it is evident that there was great variability in what the participants defined as a comfortable experience, even when constrained to 5 comfort factors. Comfort is an abstract concept and due to it being subjective, objective, personal and contextual, maybe ADS developers need to consider the need for personalisation. i.e. custom/individual comfort settings.

For low speed public transportation, operating in a constrained operating environment where all vehicles are connected and automated, then perhaps a one size fits all approach to vehicle operation/behaviour could work and a comfort algorithm can determine a baseline for comfort that the majority of passengers will be comfortable with.

However, when considering SAE L3/L4/L5 automated personal vehicles or private hire vehicles, their many shapes sizes and functions and how greatly users vary in their driving behaviour today, then surely consideration must be given to the idea of offering passengers the option to adjust parameters?

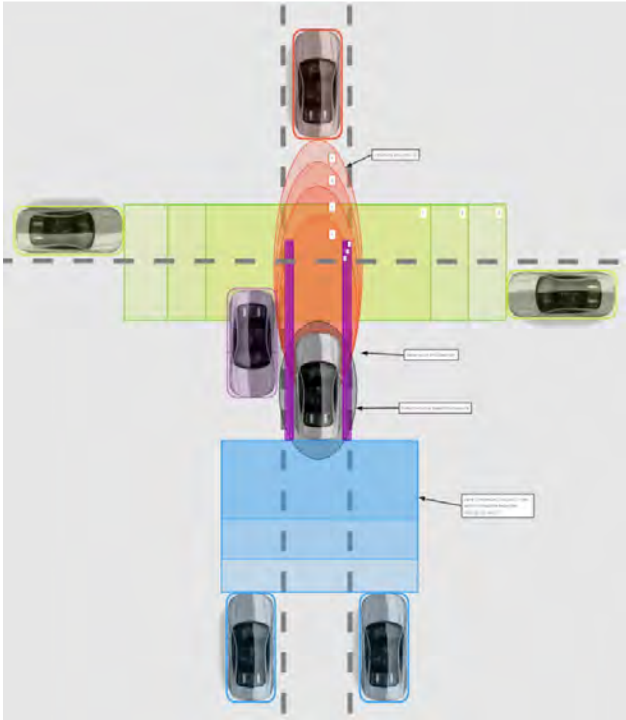


Figure 31: Shows a graphical representation of comfort that users could customise to their preferences.

Figure 31 shows a concept of how Lateral offset (gap to adjacent vehicles), Headway (Gap to vehicles ahead) and the gap to vehicles behind could be visualised and selected by occupants. Acceleration and Jerk parameters could be aligned to the vehicles core performance settings, (i.e. Eco, Comfort, Normal Sport Sport+).

Some passengers will delight in the concept of handing over full control and decision making to an automated vehicle, but others will not. Some passengers will want the system to operate only on the motorway in low traffic conditions, some will want the car to take over in busy, slow, high traffic conditions. It is hard to imagine a system that would satisfy all potential occupants/users under all works for all circumstances and for all user preferences, but it would be possible to satisfy a large proportion of people most of the time. This could be the approach for public transportation however, when considering Private owned or private hire vehicles, offering occupants the ability to influence, even marginally, how the automated vehicle operates and behaves could have a significant effect on trust, comfort and acceptance by end users.

It should be noted that the concept of personalisation described above does not diminish the research undertaken and the value of a comfort algorithm to test autonomous systems virtually/in simulations. Comfort thresholds will serve as the building blocks upon which further research can build upon and to determine how the bounds of these limits align to personality types, road conditions, environments etc.

The methods used in this research are scalable and there is the potential for researchers to build a database of evidence that the automotive industry could use in R&D. Furthermore, this evidence could serve to inform the creation of user centric standards to ensure a baseline of comfort is experienced leading to greater trust and faster adoption by the public.

9 Limitations

9.1 Technical

9.1.1 Comfort Algorithm

The comfort algorithm was developed using data from scientific peer reviewed literature. The questions that the VeriCAV project set out to answer were:

- Is the data appropriate and transferable to autonomous vehicles?
- Can an algorithm accurately reflect the perceptions and ratings of passengers?

The comfort algorithm was able to ingest the data from VeriCAV's simulation software and calculate the physics needed to compare to the comfort boundaries.

The purpose of the user trial was to verify the accuracy of the comfort rating generated by the algorithm matched the ratings given by the participants.

As a result of not being able to simulate Acceleration and Jerk in the user trial and the effects of acceleration and Jerk on the overall user experience, it has not been possible to fully verify the comfort model.

However, the scoring matrix worked satisfactorily and provided a good spread of score/ratings per 10 second segment of scenario. The Overall comfort scoring/rating system did not give account for failures of individual comfort factors within time segment (10 seconds), and therefore awarded higher than expected test results. It is therefore recommended that the overall comfort rating is defined by the lowest comfort factor rating.

9.1.2 Software

The user trial required software engineers to use the software tools provisioned through the VeriCAV project. To deliver against the needs of the user trial, they had to be modified and integrated in new and novel ways. For example, Unity and LGSVL were used to generate the simulations by reading Log files and moving the vehicles within the virtual environment every time step. To create a VR experience involved creating new assets for the vehicle interior, managing the challenges of high computational demands/refresh rates and animating aspects of the Ego/vehicles in the scene. In addition to these challenges, the need to run user trials remotely required additional steps to capture the real-time simulation via a 360video recording software product. This created additional challenges and issues to resolve. For example, the inclusion of rear view mirrors in the simulation created too high a computational load, affecting the smoothness (refresh rate) of the simulation and therefore they were removed from the Ego.

One of the biggest challenges was data logging frequency. There were data logging limits within Prescan of 20hz which was on the threshold of acceptability. This made it challenging to achieve Ego/Vehicle movements that were smooth, flicker free and natural. The end result was marginal, but acceptable for running the user trial.

Due to the open source nature of VeriCAV, it may be possible to plug in different software applications in the future, that may be more appropriate to the task of creating more realistic VR experiences.

The Apollo ADS was only used for 2 scenarios in the user trial (S4&S8). This meant that the other scenarios used an OpenScenario file to drive the movements of the Ego/vehicles. The movements and behaviours of the Ego were noticeably better and more realistic when controlled by Apollo. In a few of the OpenScenario scenarios, the Ego's cornering behaviour was visibly aggressive and appeared to rotate the Ego around a central pivot point. This was perpetuated by the road geometry and lane definitions in the OpenDrive files.

If participants had experienced the rapid rotations through a motion platform, this could have resulted in motion sickness or physical jarring. It is therefore recommended that future trials use a mature ADS. It is advisable to run the scenario through the comfort algorithm first, to determine the level of comforts experienced and determine if these are appropriate/ethical to expose to users.

There have been a lot of lessons learnt and tacit knowledge gained by VeriCAV collaboration partners. To fully leverage this knowledge, it is advisable to determine and take the next logical developmental steps to ensure the technical knowledge does not erode with time or lost through staff turnover.

9.1.3 Hardware

The Oculus Quest 2 VR headset has many benefits and capabilities. However, it could not satisfy all objectives of this project. This is not a criticism of the hardware, in fact what was delivered to participants was exemplary and met the specifications of the remote user trial plan. But, the limited field of view (110°) field of view limited participants' ability to perceive acceleration and accurately judge the velocity of the Ego. If the Quest 2 had been used with a motion platform driving simulator, it is likely results would have been improved.

The original hardware, planned for an office based user trial, was acquired prior to the work from home instructions mandated by the 3rd Covid-19 Lockdown. The hardware included:

- An RTX3090 GPU and i9 CPU gaming workstation
- A Pimax 8K VR headset with 210° FoV
- Vesaro driving simulator platform with 3" D-Box Motion actuators.

The workstation was used to generate the simulation and capture the 360 video that was loaded onto the Oculus Quest 2 VR headset. It is recommended that the intended hardware is used to conduct future user trials to verify the comfort algorithm. This will more accurately simulate the comfort factors and enable the comfort model to be verified.

9.2 Methods

9.2.1 Sensor/Scenario simulation

Prescan was used by VeriCAV's software/simulation to deliver virtual sensor information to the ADS as it navigated through the scenarios. Because the Apollo ADS was only used in 2 scenarios (S6 & S8), it is not possible to determine how effective and realistic the sensors/ADS was behaving in relation to the road geometry, the environment and other vehicles. This was not something that the comfort model/user trial had control over.

However, it is clear that to undertake a realistic user trial, that the data captured and translated into gaming engine software (Unity/LGSVL) to enable a user to experience the realistic and accurate simulation, requires a high degree of fidelity and flexibility.

For example, the data contained in the log file that Prescan generates, needs more variables to facilitate the VR simulation. For example, the inclusion of brake light application, directional indicators and steering wheel rotation would have been useful data sets that would have negated the need for manual activation/coding.

This HMI (directional indicators) was implemented manually and proved to be confidence inspiring, helping participants to anticipate directional changes. This contributed positively to their comfort / experience.

9.2.2 Vehicle behaviours and simulator realism

The behaviour of the Ego and other vehicles in the user trial/simulation exhibited behaviours that were not highly realistic. The realism ratings given by the participants averaged at 5.9/10.

As mentioned in the previous sections above, if the ADS/log file could capture the turning braking and brake light/indicator activation information, this would not only help speed up the creation of the simulation/scenarios, but it would also recreate the actual behaviour HMI/feedback mechanisms that would be controlled by the vehicle/ADS. This level of detail and fidelity contribute to passenger confidence, trust and comfort.

Details could have been improved within the simulation, including the rotation vehicle's wheels, speed based road noise simulation, higher resolution 3D models and more accurate shaders/lighting.

The positive take away, is that most of these limitations were limited by the delivery mechanism (Mobile VR headsets). Therefore, with the benefit of hindsight and using the intended hardware and greater knowledge of the software/its capabilities, generating a more realistic simulation would be a lot faster and would have generated enhanced data from the user trial.

9.2.3 User trial Experimental method

The user trial experimental delivered the goals of the trials plan and there are no limitations or operational improvements to report.

10 Next Steps

10.1 Refine comfort model

The comfort model's overall comfort rating calculation should be changed such that the overall comfort rating is determined by the lowest comfort factor rating and longer duration scenarios do not mask poor short instance performance.

10.2 Rerun the user trial simulation with motion platform

The user trial lacked motion simulation. The participants therefore lacked/did not experience what the researched deemed to be the primary components of comfort. Therefore, the trial needs to be re-run using a motion platform and wide FoV VR headset. The simulation would need to run directly from a High Performance gaming PC using a Pimax 8K VR headset (or equivalent) and motion enabled driving simulator. This will maximise the realism of the simulation and maximise the recreation/fidelity of the comfort factors under test.

10.3 Analyse the data with University partners

The user trial and software/tools built for this project collected more data than it was possible to analyse in the time available. It is recommended that a future user trial would be conducted in partnership with an Academic partner. This would provide additional domain and technical expertise and resource needed to analyse the data captured.

The software and hardware system could be duplicated to increase the number of participants involved in a study, by running the user trial at two sights simultaneously.

A further benefit of duplicating the hardware/software at multiple sights is the potential to use the same scenarios to build additive research studies. For example, there is an outstanding series of questions about CAV HMI. It is not clear what is legally required for different SAE Levels of automation and vehicle types. Multiple HMI concepts could be implemented within the Ego's cockpits and evaluated in each scenario. Alternatively, the same simulations could be used to evaluate more variables associated with vehicle proximity and roundabout merging parameters.

This could have significant impacts on the cost and time of development, ensuring that only autonomous systems that have been evaluated virtually, <u>**and experienced by humans**</u> progress to the expensive and time consuming stage of physical prototype development.

10.4 Was using VR the right choice?

VR is a new audio-visual mechanism that is still in its early stages (infancy). Connected Places Catapult have been experimenting with how, when and where to use VR to understand its potential and its limitations since 2015.

Connected Places Catapult's conclusions are that VR is not the answer to every issue. It must be applied and implemented appropriately. The cost benefit question is imperative to ask/calculate prior to embarking on a VR based simulation/solution.

Researchers selected VR as the mechanism to deliver the simulation in this project. The simulations/scenarios could have been implemented using a traditional driving simulator in a room/laboratory, but it was deemed that the technology, the user experience, the limited distractions (when wearing the headset) and the immersion could deliver significant benefits. Furthermore, the Connected Places Catapult wanted to push the boundaries.

When the decision was made to switch to a home based/remote user trial delivery VR offered significant benefits when compared to the other option of delivering videos for participants to watch on a flat screen. (although it would have been a lot easier to implement).

The Catapult has delivered on its objectives, despite the challenges created by the global pandemic and has learnt a lot of valuable lessons in the process. We hope to be able to support businesses considering the use of VR and comfort evaluation going forward. We also hope to be able to build on the work delivered to support businesses to create cutting-edge user research, within the automotive sector and beyond.

10.5 The Future of VR user trials?

Using a mobile VR headset to deliver simulated experiences, managing content, collecting data and facilitating user trials has been a fascinating experience that has generated greater knowledge about how VR can and should be used for user research. Clearly, there are only two human senses that today can be integrated into an immersive experience (Visual and Audio). However, there are new tools on the horizon that could enhance the VR experience and research potential. There are current and soon to be released VR headsets that integrate eye tracking, skin temperature, heart rate and there are additive systems such as force feedback gloves to resistance to virtual object, face/mouth tracking and olfactory simulation. These capabilities enhance the research potential and will recreate simulations that generate more useful and relevant user data.

With regards to CAVs, there is a genuine opportunity to leverage the vast numbers of Oculus VR headsets that are now in consumers hands around the globe. VR headset owners could be invited to participate in user trials/customer clinics, without the need for physical co-location. What if user trials facilitators could invite participants into a simulated vehicle and sit next to them (virtually) as the scenarios unfold, collecting heart rate data, brain activity, stress levels etc.? The only alternative today to this application of VR, is physical testing, which is very expensive, especially if developing a product that will be sold across the world.

Another concept that has merit in relation to CAV comfort, is one of Human in the loop testing and using AI to process participants driving behaviour in a simulator. For example, a participant could sit in a driving simulator wearing a VR headset and be exposed to an ADS driven scenario. The participant could then, using a steering wheel and pedals, drive the same scenario. AI could learn what the participant did differently to the ADS and with enough data sets, could suggest/implement improvements to the ADS's behaviour.

In summary, businesses and academia have an opportunity to leverage these new cutting-edge software and hardware technologies to support the creation of human centric automated solutions. The Connected Places Catapult is here to help, advise and would welcome opportunities to build on the lessons learnt and experiences enabled through our involvement in the VeriCAV project.

10.6 Final Remarks

The data collected in the user trial showed that the comfort algorithm has a strong foundation on which occupant comfort can be predicted, despite the limitations of the hardware used.

The Comfort Algorithm requires further validation testing using vehicle dynamics simulation and using a wide-angle Field of View VR headset to replicate the experience necessary to evaluate the comfort algorithm’s 5 comfort factors.

Remote user trials can be facilitated through VR effectively and have cost and time saving benefits. However, the technology must be implemented effectively to deliver the necessary research outcomes.

The process by which scenarios were generated and used for Human/participants verification has been developed so that scenarios and interactivity can be iterated quickly. This has significant implications for design and development of new automotive concepts, saving time and cost, but it also opens up opportunities to a wider range of professions, allowing greater freedom to experiment and refine concepts prior to real world construction and testing.

11 Outcomes and Innovation

- Faster than real time assessment and prediction of basic occupant comfort is possible using the VeriCAV Comfort Test oracle.

- Using the latest Mobile VR technology, Connected Places Catapult were able to deliver user trials remotely and have developed methodologies and practical measures that can be replicated using off the shelf hardware.

- The simulation capability developed during the VeriCAV project can be used to support future research into the Human Factors of Autonomous Vehicles. These include assessment of vehicle behaviour, user interface evaluation (HMI), trust, comfort, and user experience.

- The VeriCAV Autonomous Vehicle Simulator can be easily upgraded to work with other simulation software and can use trajectory and vehicle status files from Automated Drive Systems.

- Connected Places Catapult have developed practical and Covid safe methodologies to enable Health and Safety compliant user trials.

References

1. Bae, J. Moon, and J. Seo, “Toward a Comfortable Driving Experience for a Self-Driving Shuttle Bus,” pp. No. 8, Electronics (2019)

2. Bellem, H., Schönenberg, T., Krems, J. F., & Schrauf, M. (2016). Objective metrics of comfort: developing a driving style for highly automated vehicles. Transportation Research Part F: Traffic Psychology and Behaviour, 41, 45–54.

3. Diels, C., Erol, T., Kukova, M., Wasser, J., Cieslak, M., Payre, W., Miglani, A., Mansfield, N.J., Hodder, S.G. & Bos, J. (2017). Designing for comfort in shared and automated vehicles (SAV): a conceptual framework. Presented at the 1st International Comfort Congress (ICC2017), Salerno, Italy (2017)

4. Felix Wilhelm Siebert, Michael Oehl, Hans-Rüdiger Pfister ‘The influence of time headway on subjective driver statesin adaptive cruise control’, 2014.

5. Patrick Rossner & Angelika C. Bullinger ‘Does driving experience matter? Influence of trajectory behaviour ondrivers’trust, acceptance and perceived safety in automated driving’, 2019.

6. Gudrun M.I. Voß, Caroline M. Keck, Maximilian Schwalm, ‘Investigation of drivers’ thresholds of a subjectively accepteddriving performance with a focus on automated driving’, 2018.

Bibliography

1. Bae, J. Moon, and J. Seo, “Toward a Comfortable Driving Experience for a Self-Driving Shuttle Bus,” pp. No. 8, Electronics (2019)
2. Basu C, Yang Q, Hungerman D, Sinahal M and Draqan AD (2017). Do you want your autonomous car to drive like you? In 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 417-425. IEEE.
3. Bellem, H., Thiel, B., Schrauf, M. & Krems, J.F. (2018). Comfort in Automated Driving: An Analysis of Preferences for Different Automated Driving Styles and Their Dependence on Personality Traits. Transportation Research Part F: Traffic Psychology and Behaviour, 55, 90-100.
4. Ekman, F., Johansson, M., Bligård, L.-O., Karlsson, M., & Strömberg, H. (2019). Exploring automated vehicle driving styles as a source of trust information. Transportation Research Part F: Traffic Psychology and Behaviour, 65, 268-279.
5. Elbanhawi, M.; Simic, M.; Jazar, R. (2015). In the Passenger Seat: Investigating Ride Comfort Measures in Autonomous Cars. IEEE Intelligent Transport System Magazine, 7, 4–17.
6. Hartwich, F.; Beggiato, M.; Krems, J.F. (2018). Driving comfort, enjoyment and acceptance of automated driving–effects of drivers’ age and driving style familiarity. Ergonomics, 61, 1017–1032
7. Rossner, P. and Bullinger, A.C. (2019). Do You Shift or Not? Influence of Trajectory Behaviour on Perceived Safety During Automated Driving on Rural Roads. In: Krömker H. (eds) HCI in

Mobility, Transport and Automotive systems. HCII 2019. Lecture Notes in Computer Science. 11596

8. Voss, G.M.I., Keck, C.M. & Schwalm, M. (2018). Investigation of drivers' thresholds of a subjectively accepted driving performance with a focus on automated driving. Transport Research Part F: Traffic Psychology and Behaviour. 56, 280-292
9. Yusof, N.M.; Karjanto, J.; Terken, J.; Delbressine, F.; Hassan, M.Z.; Rauterberg, M. (2016). The exploration of autonomous vehicle driving styles: Preferred longitudinal, lateral, and vertical accelerations. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ann Arbor, MI, USA, 24–26; pp. 245–252.
10. Kaur, K., & Ramprasad, G. (2018). Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. Journal of Engineering and Technology Management, 48(2), 87-96.
11. Nordhoff, S.; de Winter, J.; Madigan, R.; Merat, N.; van Arem, B.; Happee, R. (2018). User acceptance of automated shuttles in Berlin-Schöneberg: A questionnaire study. Transp. Res. Part Traffic Psychol. Behav. 2018, 58, 843–854.
12. Nordhoff, S.; de Winter, J.; Payre, W.; van Arem, B.; Happee, R. What impressions do users have after a ride in an automated shuttle? An interview study. Transp. Res. Part Traffic Psychol. Behav. 2019, 63, 252–269.
13. Wasser, J., Diels, C. & Tovey, M. (2017). Driverless Pods: From Technology Demonstrators to Desirable Mobility Solutions. Advances in Intelligent Systems and Computing.
14. Experiences of Advanced Driver Assistance Systems amongst Older Drivers: An Evidence review for the DfT
15. Ferdinand Schockenhoff, Hannes Nehse, Markus Lienkamp. Maneuver-Based Objectification of User Comfort Affecting Aspects of Driving Style of Autonomous Vehicle Concepts. *Applied Sciences* **2020,** *10* (11), 3946.
16. Y. Shi, M. Bordegoni, G. Caruso (2020) User Studies by Driving Simulators in the Era of Automated Vehicle, Computer-Aided Design and Applications 18(1), p. 211-226.
17. Feng, Y., Pickering, S., Chappell, E., Iravani, P., & Brace, C. (2019). A Support Vector Clustering Based Approach for Driving Style Classification. International Journal of Machine Learning and Computing, 9(3), 344–350.
18. Bernhard, C., Oberfeld, D., Hoffmann, C., Weismüller, D., & Hecht, H. (2020). User acceptance of automated public transport. Transportation Research Part F: Traffic Psychology and Behaviour. 70. 109-123.
19. Oliveira, L., Proctor, K., Burns, C.G., and Birrell, S. (2019) *Driving style : how should an automated vehicle behave?* Information, 10 (6). 219.
20. Schwarting, W., Pierson, A., Alonso-Mora, J., Karaman, S., & Rus, D. (2019). Social behavior for autonomous vehicles. Proceedings of the National Academy of Sciences, 116(50), 24972–24978. https://doi.org/10.1073/pnas.1820676116
21. Markkula, G., Madigan, R., Nathanael, D., Portouli, E., Lee, Y.M., Dietrich, A., Billington, J., Schieben, A., and Merat, N. (2020) Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic. Theoretical Issues in Ergonomics Science
22. Li, G., Zhu, F., Qu, X., Cheng, B., Li, S., & Green, P. (2019). Driving Style Classification Based on Driving Operational Pictures. IEEE Access, 7, 90180–90189.
23. Karjanto, J., Yusof, N., Terken, J. M. B., Delbressine, F. L. M., Hassan, M. Z. B., Rauterberg, G. W. M. (2017). Simulating autonomous driving styles: accelerations for three road profiles. MATEC Web of Conferences, 90, 1- 16.

24. Karjanto, J., Md. Yusof, N., Wang, C., Terken, J., Delbressine, F., & Rauterberg, M. (2018). The effect of peripheral visual feedforward system in enhancing situation awareness and mitigating motion sickness in fully automated driving. Transportation Research Part F: Traffic Psychology and Behaviour, 58, 678–692
25. Karjanto, J., Md. Yusof, N., Terken, J., Delbressine, F., Rauterberg, M., & Hassan, M. Z. (2018). Development of on-road automated vehicle simulator for motion sickness studies (Manuscript in preparation).
26. Human Factors Design Guidance for Level 2 And Level 3 Automated Driving Concepts Link
27. Matthew C. Best (2018): Real-time characterisation of driver steering behaviour, Vehicle System Dynamics
28. Mühlbacher, D., Tomzig, M., Reinmüller, K., & Rittger, L. (2020). Methodological Considerations Concerning Motion Sickness Investigations during Automated Driving. Information, 11(5), 265.
29. Itkonen, T. H., Lehtonen, E., & Selpi. (2020). Characterisation of motorway driving style using naturalistic driving data. Transportation Research Part F: Traffic Psychology and Behaviour, 69, 72–79
30. Itkonen, T. H., Pekkanen, J., Lappi, O., Kosonen, I., Luttinen, T., & Summala, H. (2017). Trade-off between jerk and time headway as an indicator of driving style. PLOS ONE, 12(10), e0185856
31. Zhang, K., Zhang, X., Deng, W., Yu, F., Shang, S., & Guo, M. (2020). Path planning algorithm of the intelligent vehicle considering the passenger feelings. IOP Conference Series: Materials Science and Engineering, 768, 042002.
32. Salter, S., Diels, C., Herriotts, P., Kanarachos, S., & Thake, D. (2019). Model to predict motion sickness within autonomous vehicles. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 234(5), 1330–1345
33. Ulahannan, A., Cain, R., Thompson, S., Skrypchuk, L., Mouzakitis, A., Jennings, P., & Birrell, S. (2020). User expectations of partial driving automation capabilities and their effect on information design preferences in the vehicle. Applied Ergonomics, 82, 102969
34. Ruijten, P. A. M., Terken, J. M. B., & Chandramouli, S. N. (2018). Enhancing trust in autonomous vehicles through intelligent user interfaces that mimic human behavior. *Multimodal Technologies and Interaction*, *2*(4), [62]
35. Oliveira, L., Burns, C., Luton, J., Iyer, S., & Birrell, S. (2020). The influence of system transparency on trust: Evaluating interfaces in a highly automated vehicle. Transportation Research Part F: Traffic Psychology and Behaviour, 72, 280–296.
36. Ruijten, P., Terken, J., & Chandramouli, S. (2018). Enhancing Trust in Autonomous Vehicles through Intelligent User Interfaces That Mimic Human Behavior. Multimodal Technologies and Interaction, 2(4), 62.
37. Oliveira, L., Burns, C., Luton, J., Iyer, S., & Birrell, S. (2020). The influence of system transparency on trust: Evaluating interfaces in a highly automated vehicle. Transportation Research Part F: Traffic Psychology and Behaviour, 72, 280–296.
38. Beggiato, M., Hartwich, F., & Krems, J. (2018). Using Smartbands, Pupillometry and Body Motion to Detect Discomfort in Automated Driving. Frontiers in Human Neuroscience, 12.

Acknowledgements

The author would like to acknowledge the following:

- Connected Places Catapult's software development team that managed to create the software tools and simulation despite all the challenges of highly complex system integration/compatibility challenges, amidst the Covid pandemic.
- CleanBox: for providing the vital equipment needed to sterilise the VR headsets, helping to keep our staff/participants safe.
- Jack Bebbington and Mark Wade for their support and enthusiasm to drive forward the research and enable the researchers to experiment with new ways of using VR for user experience/user testing.
- Thomas Levermore (Connected Places Catapult), for working tirelessly with the researchers to troubleshoot and develop the comfort algorithm and always offering a helping hand to resolve technical challenges.
- Victoria Phillips (Computacentre), for providing the Oculus Quest 2 enterprise headsets at very short notice.

Appendix A

Comfort Data

How to read the Comfort Data

The user trial collected data from participants following an exposure to each scenario. This section presents the outputs from the participant's subjective/perceived comfort ratings and compares them to the comfort rating attributed by the Comfort Algorithm.

Participant data collection:

Following each exposure to a Scenario, participants answered a questionnaire which was broken down into 3 parts.

1. They described the scenario they experienced (verbatim)
2. They rated each comfort factor they experienced during the scenario
3. They rated the overall comfort for the scenario.

To help the reader interpret the data in this section, each chart and its purpose is described below.

Rating of each comfort factor:

The comfort factors defined in the comfort algorithm, needed to be translated into a language that was more natural and intuitive to the participants. The comfort factors were therefore described as follows:

Comfort Factor	**Participant terminology**
Acceleration and Jerk	'Acceleration' and 'Deceleration'.
Headway	'Gap to the vehicle ahead'
Lateral offset	'the proximity of cars passing to the side of your car'
	'your vehicle's lane positioning'
Turning at T-junctions and joining roundabouts (Gap)	'how comfortable were you when your car pulled out onto the roundabout or pulled out of the T-Junction'.

Figure 32: Shows the terminology used in the Comfort Algorithm and how this was translated into familiar language for the participants.

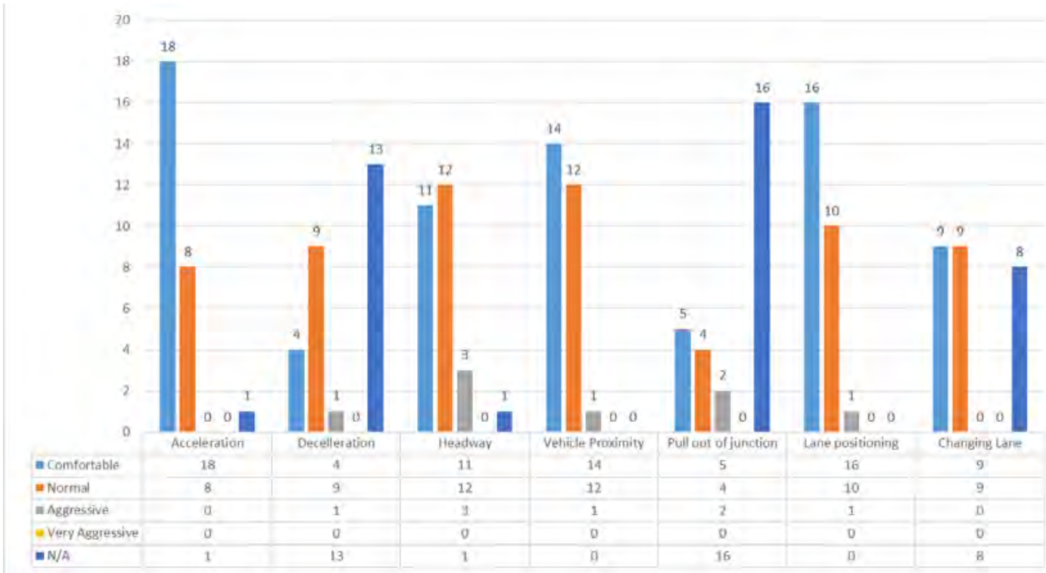
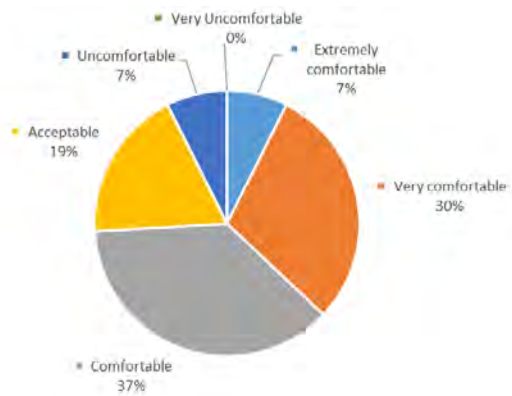


Figure 33: An example of the chart used to summarise participant comfort ratings for the individual comfort factors.

Overall comfort rating for the Scenario

Following the individual comfort factor rating, participants were asked to assign an Overall Comfort Rating for the Scenario. This is shown using the table and pie chart below. Asking the participants to provide an overall comfort rating provided the researchers the ability to identify if individual negative comfort factor ratings affected participant's overall comfort rating for the scenario.



	Overall Comfort Rating	Percentage
2	Extremely comfortable	7%
8	Very comfortable	30%
10	Comfortable	37%
5	Acceptable	19%
2	Uncomfortable	7%
0	Very Uncomfortable	0%
27		

Figure 34: Overall comfort rating data

The above table/pie chart shows the number of participants assigning each comfort rating and the percentage split of responses.

For the purposes of analysis, the assumption is that participant's responses that are 'Comfortable', 'Very Comfortable' and 'Extremely Comfortable' are ratings that are positive. Ratings of 'Acceptable' or below and considered to be negative and prompted a review of subjective feedback to identify common themes that may have resulted in negative feedback within the participants.

THE COMFORT ALGORITHM / ORACLE COMFORT SCORING PROCESS

There are 4 comfort factors measured by the Comfort Algorithm:

- Combined acceleration (Lateral and longitudinal)
- Combined Jerk (Lateral and Longitudinal)
- Headway to vehicle ahead (same lane)
- Lateral offset (distance to adjacent vehicles)
- Gap (time in seconds between two cars for the Ego to pull into).

Within a scenario, the algorithm deducts points from a starting score maximum of 100. For every second that thresholds are exceeded, points are deducted relative to the magnitude of and duration. See sample time scoring matrix below:

Sample Time/ negative Scoring matrix:

	Sample time	**Starting Score**
	10s	**100pts**
Comfort levels	**Negative** Scoring / sec	Secs of exposure to result in a fail
Comfortable	0	n/a
Normal driving	1	40
Aggressive/ uncomfortable	10	4
Extremely Aggressive/ very uncomfortable	20	2

Figure 35 Shows the matrix used by the comfort algorithm to deduct points for breaching comfort thresholds.:

As the scenarios for the trial were under 30 seconds in duration, a journey that was perceived as 'normal' would not result in a fail.

Comfort Grading:

The comfort grading chart below is used to provide a comfort grade for the individual comfort factors. The same scoring system is used to grade the overall Scenario.

COMFORT Grade/scoring chart

Comfort Grades	**Comfort score Min**	**Comfort Score Max**	**Description**

A*	90	100	Extremely comfortable
A	80	89	Very comfortable
B	75	79	Comfortable
C	65	69	Acceptable
D	60	64	Poor
F	0	59	Fail / uncomfortable

Figure 36: Shows the Comfort Grading and Scoring look up chart.

The Comfort Oracle Scenario table:

The following table shows an example output of how the comfort algorithm scored for the individual comfort factors and the overall comfort rating for the scenario.

	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars
Seconds	**CF1A**	**CF1B**	**CF2**	**CF3**	**CF4 : Not measured!**
10	0	15.05	100	100	100
20	99.55	99.75	100	100	100
30	100	100	100	100	100
40	100	100	100	90.8	100
50	100	100	100	18.7	100
Averages	79.91	82.96	100	81.9	100
Comfort Grade	B	A	A*	A	N/A
Pass/Fail	FAIL	FAIL	PASS	FAIL	N/A
Overall comfort score	86.1925				
Overall Comfort PASS/FAIL	**FAIL**	A			

Figure 37: Shows the raw data output from the Comfort Algorithm

*Note: Any individual comfort factor scoring below 60 points (highlighted in red), in a 10 second period of time, will result in an **overall comfort test fail**.*

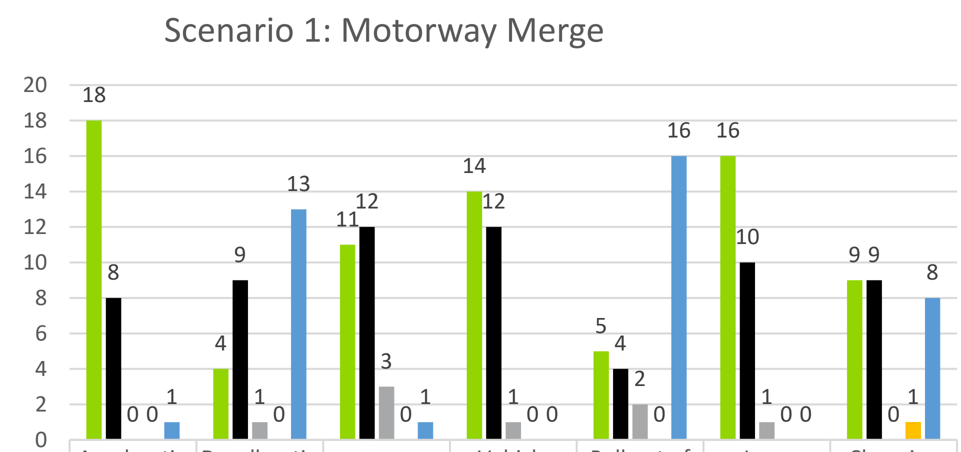
Note: The Oracle/VeriCAV software was not able to reliably calculate the scores for 'Gap between cars' (pulling out from a junction/onto a roundabout). For applicable scenarios (S4&S8), the gap was calculated manually and scoring deductions applied accordingly.

COMFORT RESULTS

SCENARIO 1: Motorway Merge

Watch the Scenario video and the Comfort Algorithm side by side here: [LINK]

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	18	4	11	14	5	16	9
Normal	8	9	12	12	4	10	9
Aggressive	0	1	3	1	2	1	0
Extremely Aggressive	0	0	0	0	0	0	1
N/A	1	13	1	0	16	0	8

Participant Comfort factor results:

Only 1 Participant scored the deceleration negatively in the scenario. This indicates that participants may not have been able to detect and quantify acceleration using the simulation display equipment (VR headset). This indicates that a combination of factors such as (but not limited to) the lack of stimulus/haptics to simulate acceleration and/or the limited field of view of the headset, may be responsible for this result.

The following comfort factor received an **extremely aggressive** rating:

Changing Lane = 1 (3.7%)

The following comfort factors received **aggressive** ratings:

Headway = 3 (11%)

Proximity = 1 (3.7%)

Pulling out of junction = 2 (7.4%)

Lane positioning = 1 (3.7%)

Verbatim Analysis

This scenario involved merging onto a motorway, and the most common theme from qualitative comments were that there were no wing mirrors and so participants couldn't see behind them. Some participants voiced that this felt limiting and they tried to move their head more to check the inside lane, as they 'couldn't see cars coming/overtaking till they were next to you'. This lack of situational awareness was cited as the reason that some participants felt uncomfortable, and there was a comment made that 'it used indicators, but wanted to verify myself' and 'having a mirror or screen would help comfort'. There were some comments stating that the ego vehicle 'got close to the car ahead' and some participants reported pulling the trigger as a result of this.

Raw data from the comfort algorithm:

	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars
Seconds	**CF1A**	**CF1B**	**CF2**	**CF3**	**CF4 : Not measured!**
10	0	15.05	100	100	100
20	99.55	99.75	100	100	100
30	100	100	100	100	100
40	100	100	100	90.8	100
50	100	100	100	18.7	100
Averages	79.91	82.96	100	81.9	100
Comfort Grade	B	A	A*	A	N/A
Pass/Fail	FAIL	FAIL	PASS	FAIL	N/A
Overall comfort score	86.1925				
Overall Comfort PASS/FAIL	**FAIL**	A			

Scenario 1 Data analysis:

The Comfort algorithm results: 3 comfort factors failures.

1. **Acceleration:** At the beginning of the scenario, the algorithm deemed this as too aggressive (score of 0). Analysis shows that the ego vehicle accelerated approximately from 0-60mph in 7 seconds.
2. **Jerk**: In addition to this, the acceleration occurred whilst the ego vehicle was going round a curve on the slip lane of the motorway (score of 15.05).
3. **Headway**: At the end of the scenario, the ego's headway (gap to the car ahead) was identified as too small (score of 18.7).

Overall comfort scoring:

	Participant Results		Algorithm Results		Scoring Matrix		
No. of	**Percentage**	**Overall Comfort Rating**	**Comfort Score**	**Pass/ Fail**	**Algorithm Grades**	**Algorithm Ratings**	**Overall Ave Score**
2	7%	Extremely comfortable			A*	Extremely comfortable	>90
8	30%	Very comfortable			A	Very comfortable	>80
10	**37%**	**Comfortable**			B	Comfortable	>75
5	19%	Acceptable			C	Acceptable	>65
2	7%	Uncomfortable			D	Poor	>60
0	0%	Very Uncomfortable		**FAIL**	F	Fail / uncomfortable	<60
27							

Participants:

25 (93%) of the participants, rated the scenario experience as Comfortable, Very comfortable or Extremely comfortable.

2 (7%) of the participants rated the Scenario as Acceptable, Uncomfortable or Very Uncomfortable.

Comfort Algorithm:

The Comfort algorithm generated a FAIL.

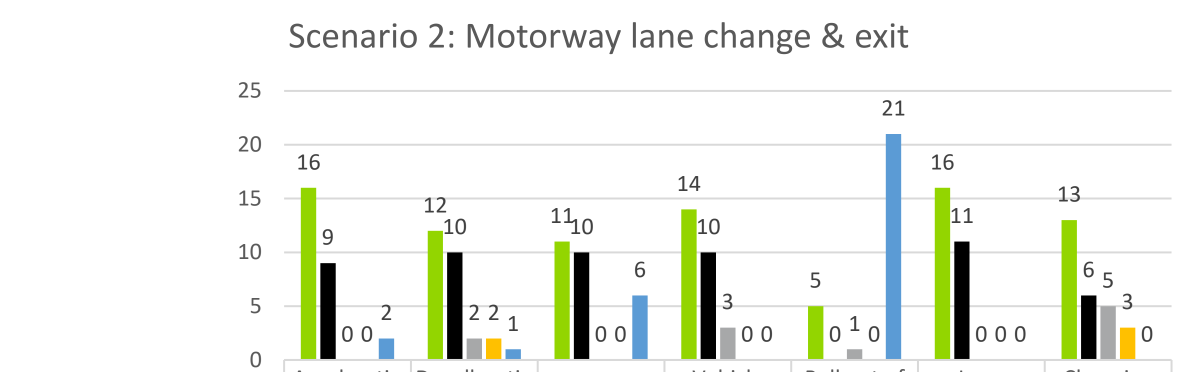
Alignment:

No alignment between participant's scores and the comfort oracle.

SCENARIO 2: Motorway Lane Change and Exit

Watch the Scenario video, Comfort Algorithm and Scoring here: LINK

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	16	12	11	14	5	16	13
Normal	9	10	10	10	0	11	6
Aggressive	0	2	0	3	1	0	5
Extremely Aggressive	0	2	0	0	0	0	3
N/A	2	1	6	0	21	0	0

Participant overall comfort ratings for Scenario 2: Motorway lane change and exit

Participant Comfort factor results:

Participants ratings were very positive for the comfort factors scoring Comfortable or Normal. There were a few participants that rated comfort factors as Aggressive of Extremely Aggressive:

Aggressive:

- Deceleration= 2
- Vehicle proximity= 3
- Pulling out of junction= 1
- Changing Lane = 5

Extremely Aggressive:

- Deceleration = 2
- Changing Lane = 3

Verbatim Analysis

This was changing lanes on a motorway and then taking a motorway exit. There was a mixed spread of themes with no clear prevalent theme. Participants largely summarised what occurred during the scenario, commenting that there was 'no change in acceleration or deceleration as it pulled off' and that they 'felt comfortable at all times, didn't pull the trigger'.

Raw data from the comfort algorithm:

Comfort Algorithm Scores: Scenario 2					
	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars
Seconds	CF1A	CF1B	CF2	CF3	CF4 : Not measured
10	100	100	100	100	100
20	100	100	100	100	100
30	100	100	100	100	100
40	94.4	100	100	100	100
50	97.45	100	100	100	100
Averages	98.37	100	100	100	100
Comfort Grade	A*	A*	A*	A*	N/A
Pass/Fail	PASS	PASS	PASS	PASS	N/A
Overall comfort score	99.5925				
Overall Comfort PASS/FAIL	**PASS**	A*			

Scenario 2 Data analysis:

The Comfort algorithm results:

*The comfort algorithm reported 0 comfort factor failures with no scores less than 94.4 = A**

Overall comfort scoring

	Participant Results			Algorithm Results		Scoring Matrix		
No. of	**Percentage**	**Overall Comfort Rating**		**Comfort Score**	**Pass/ Fail**	**Algorithm Grades**	**Algorithm Ratings**	**Overall Ave Score**
3	11%	Extremely comfortable		**A***	**PASS**	A*	Extremely comfortable	>90
10	**37%**	**Very comfortable**				A	Very comfortable	>80
6	22%	Comfortable				B	Comfortable	>75
5	19%	Acceptable				C	Acceptable	>65
3	11%	Uncomfortable				D	Poor	>60
0	0%	Very Uncomfortable				F	Fail / uncomfortable	<60
27								

Participants:

89% of the participants, rated the scenario experience as **Acceptable**, **Comfortable**, **Very Comfortable** or **Extremely comfortable**.

11% of the participants rated the Scenario as, **Uncomfortable**.

Comfort Algorithm:

The Comfort algorithm generated a PASS and awarded an **Extremely Comfortable** rating with a score of A*.

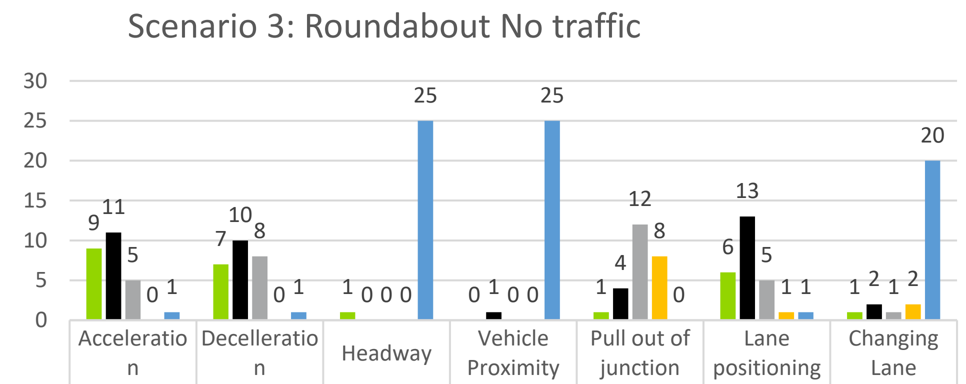
Alignment:

Positive Alignment: The majority of the participants **48%** rated the overall experience as **Very Comfortable** (**A**) **37%** or rated it as **Extremely Comfortable (A*)11%**.

SCENARIO 3: Roundabout No Traffic

Watch the Scenario video, Comfort Algorithm and Scoring here: [LINK]

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	9	7	1	0	1	6	1
Normal	11	10	0	1	4	13	2
Aggressive	5	8	0	0	12	5	1
Extremely Aggressive	0	0	0	0	8	1	2
N/A	1	1	25	25	0	1	20

Participant overall comfort ratings for Scenario 3: Roundabout no traffic

Participant Comfort factor results:

There were a large number of participants that recorded **Aggressive** comfort ratings:

- Acceleration = 5
- Deceleration= 8
- Pulling out of junction= 12
- Lane Positioning = 5
- Lane Change = 1

There were a large number of participants that recorded **Extremely Aggressive** comfort ratings:

- Pull out of junction = 8
- Lane Positioning = 1
- Changing Lane = 2

Verbatim analysis:

Here the AV joined a roundabout and then exited on the second turn-off with no other cars present. Participants' most prevalent comments were that they were uncomfortable with the joining of the roundabout (or turning if they perceived it as a T junction). They perceived the turns as 'very sharp' and 'very robotic and spinning on the spot'. Some participants also expressed surprise that it didn't come to a stop at the junction and that they felt like it 'went almost on the other side of the road'.

Raw data from the comfort algorithm:

Comfort Algorithm Scores: Scenario 3					
	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars
Seconds	CF1A	CF1B	CF2	CF3	CF4 : Not measured
10	0	75.5	100	100	100
20	49.5	59.7	100	100	100
30	90	88.2	100	100	100
40	82.35	69.4	100	100	100
50	98.35	100	100	100	100
Averages	64.04	78.56	100	100	100
Comfort Grade	D	B	A*	A*	N/A
Pass/Fail	FAIL	PASS	PASS	PASS	N/A
Overall comfort score	85.65				
Overall Comfort PASS/FAIL	**FAIL**	A			

Scenario 3 Data analysis:

The Comfort algorithm results:

The comfort algorithm reported 2 comfort factor failures:

1. Acceleration
2. Jerk

Both of these failures occurred when the Ego joined the roundabout. There were negative deductions as the ego travelled round the roundabout, but it did not fail any subsequent 10 second segments thereafter. This indicates that the of the braking and turning characteristics of the ADS should be adjusted to improve comfort.

Overall comfort scoring

	Participant Results			Algorithm Results			Scoring Matrix		
No. of	Percentage	Overall Comfort Rating		Comfort Score	Pass/ Fail		Algorithm Grades	Algorithm Ratings	Overall Ave Score
0	0%	Extremely comfortable					A*	Extremely comfortable	>90
2	8%	Very comfortable					A	Very comfortable	>80
9	**35%**	**Comfortable**					B	Comfortable	>75
6	23%	Acceptable					C	Acceptable	>65
8	31%	Uncomfortable					D	Poor	>60
1	4%	Very Uncomfortable			**FAIL**		F	Fail / uncomfortable	<60
26									

Participants:

65% of the participants, rated the scenario experience as Comfortable, Very comfortable or Extremely comfortable.

35% of the participants rated the Scenario as acceptable, Uncomfortable or Very Uncomfortable.

Comfort Algorithm:

The Comfort algorithm generated a FAIL*.

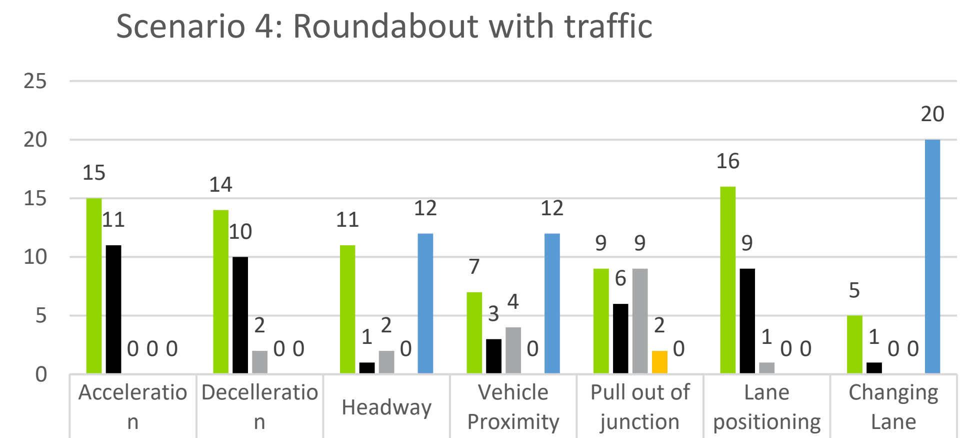
Alignment:

No Alignment: Although 35% of the participants negatively scored the experience, there were still 66% who rated the ADS as acceptable or better. There is a strong likelihood that this is due to a lack of accelerative/G forces being simulated.

SCENARIO 4: Roundabout with Traffic

To watch a video of the Scenario and the graphs Comfort Algorithm graphs: [LINK]

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	15	14	11	7	9	16	5
Normal	11	10	1	3	6	9	1
Aggressive	0	2	2	4	9	1	0
Extremely Aggressive	0	0	0	0	2	0	0
N/A	0	0	12	12	0	0	20

Participant overall comfort ratings for Scenario 4: Roundabout with traffic

Participant Comfort factor results:

In this roundabout scenario, the deceleration, acceleration and turning was much smoother than Scenario 3. The rating given by participants for these comfort factors was much improved as a result.

The following No. of participants recorded **Aggressive** comfort ratings:

- Deceleration= 2
- Headway = 2
- Vehicle Proximity = 4
- Pulling out of junction = 9
- Lane Positioning = 1

The following No. of participants recorded **Extremely Aggressive** comfort ratings:

- Pull out of Junction = 2

Given the proximity of the vehicle approaching on the roundabout, it was surprising that more participants did not rate the ‘Pulling out of junction’ comfort rating more negatively

Verbatim analysis:

As with scenario 3 this involved joining a roundabout and exiting, though there were also 2 cars on the roundabout. Participants stated they were happy with the speed that the ego vehicle travelled at, but they expected it to stop and give way to the oncoming vehicle and it didn't. They stated that 'I would have waited for the car to pass before I would have pulled out', and 'not sure I would have gone ahead'. Some recognised that it was a very similar scenario to the previous one but they felt the turning was 'smoother' and that 'it didn't come out as far into the road as went round the corner'.

Raw data from the comfort algorithm:

Comfort Algorithm Scores: Scenario 4

	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars	
Seconds	CF1A	CF1B	CF2	CF3	CF4: Measured Manually	3 secong gap to car on roundabout, exposure of 4 seconds therefore: 100 - (20x4) = 20
10	64.1	81.5	100	100	20	
20	94.1	100	100	100	100	
30	90	98.55	100	100	100	
40	72.3	96	100	100	100	
50	77.85	100	100	100	100	
Averages	79.67	95.21	100	100	84	
Comfort Grade	B	A*	A*	A*	A	
Pass/Fail	PASS	PASS	PASS	PASS	FAIL	
Overall comfort score			93.72			
Overall Comfort PASS/FAIL	**FAIL**	A*				

Scenario 4 Data analysis:

The Comfort algorithm results:

The comfort algorithm reported 1 comfort factor failures:

1. Gap

This failure was due to the gap between the Ego car and the car already on the roundabout as it joined the roundabout. The gap was measured manually due to the comfort algorithm not working reliably for the gap comfort factor. The gap in seconds was measured as 3 seconds, which is rated as 'extremely aggressive'.

Overall comfort scoring

	Participant Results		Algorithm Results		Scoring Matrix		
No. of	**Percentage**	**Overall Comfort Rating**	**Comfort Score**	**Pass/ Fail**	**Algorithm Grades**	**Algorithm Ratings**	**Overall Ave Score**
1	4%	Extremely comfortable			A*	Extremely comfortable	>90
8	31%	Very comfortable			A	Very comfortable	>80
7	27%	Comfortable			B	Comfortable	>75
8	31%	Acceptable			C	Acceptable	>65
2	8%	Uncomfortable			D	Poor	>60
0	0%	Very Uncomfortable		**FAIL**	F	Fail / uncomfortable	<60
26							

Participants:

92% of the participants, rated the scenario experience as Acceptable, Comfortable, Very comfortable or Extremely comfortable.

8% of the participants rated the Scenario as, Uncomfortable or Very Uncomfortable.

Comfort Algorithm:

The Comfort algorithm generated a FAIL*.

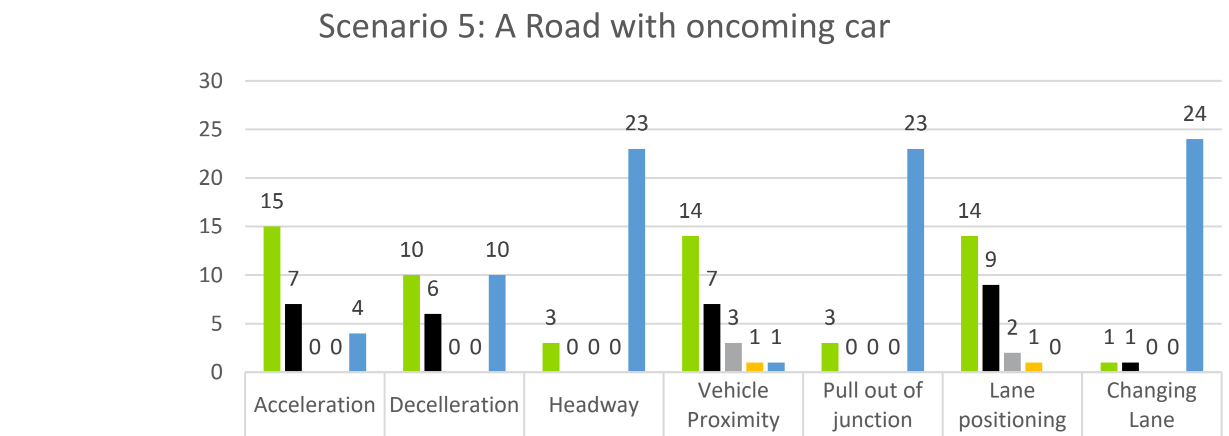
Alignment:

No Alignment: Although 38% of the participants rated the pulling out at junction as Aggressive or Extremely aggressive, this did not affect their overall scenario rating where 92% of the participants stated that the scenario was Acceptable, Comfortable, Very Comfortable or Extremely comfortable.

SCENARIO 5: A Road, Oncoming Car

Watch the Scenario video, Comfort Algorithm and Scoring here: [LINK]

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	15	10	3	14	3	14	1
Normal	7	6	0	7	0	9	1
Aggressive	0	0	0	3	0	2	0
Extremely Aggressive	0	0	0	1	0	1	0
N/A	4	10	23	1	23	0	24

Participant overall comfort ratings for Scenario 5: A Road, Oncoming car

Participant Comfort factor results:

The following No. of participants recorded **Aggressive** comfort ratings:

- Vehicle Proximity = 3
- Lane Positioning = 2

The following No. of participants recorded **Extremely Aggressive** comfort ratings:

- Vehicle Proximity = 1
- Lane Positioning = 1

Verbatim analysis:

This scenario involved driving on an A road and passing an oncoming car, after which participants summarised that they were happy with the speed of travel. Participants observed that there was a car coming the other way but stated that overall, the scenario was ‘uneventful’ and a ‘nice comfortable easy drive’. A small number of participants found the second turning ‘sharp’ and one pulled the trigger as a result of this.

Raw data from the comfort algorithm:

Comfort Algorithm Scores: Scenario 5

	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars
Seconds	CF1A	CF1B	CF2	CF3	CF4: Not measured
10	93.5	100	100	100	100
20	97.1	100	100	100	100
30	98.65	100	96	100	100
40	98.15	100	100	100	100
Averages	96.9	100	99	100	N/A
Comfort Grade	A*	A*	A*	A*	N/A
Pass/Fail	PASS	PASS	PASS	PASS	N/A
Overall comfort score	99.0				
Overall Comfort PASS/FAIL	**PASS**	A*			

Scenario 5 Data analysis:

The Comfort algorithm results:

The comfort algorithm reported 0 comfort factor failures

Overall comfort scoring

	Participant Results		Algorithm Results		Scoring Matrix		
No. of	**Percentage**	**Overall Comfort Rating**	**Comfort Score**	**Pass/ Fail**	**Algorithm Grades**	**Algorithm Ratings**	**Overall Ave Score**
4	15%	Extremely comfortable	A*	**PASS**	A*	Extremely comfortable	>90
10	**38%**	**Very comfortable**			A	Very comfortable	>80
9	35%	Comfortable			B	Comfortable	>75
2	8%	Acceptable			C	Acceptable	>65
1	4%	Uncomfortable			D	Poor	>60
0	0%	Very Uncomfortable			F	Fail / uncomfortable	<60
26							

Participants:

96% of the participants, rated the scenario experience as Acceptable, Comfortable, Very comfortable or Extremely comfortable.

1 (4%) of the participants rated the Scenario as, Uncomfortable.

Comfort Algorithm:

The Comfort algorithm generated a PASS with an A* grade.

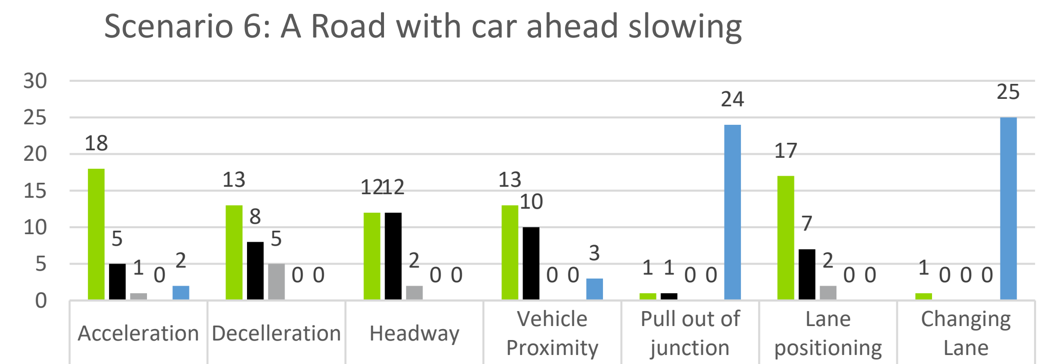
Alignment:

There was a good relationship between the participant's ratings and that of the Algorithm, although participants did not rate the comfort as highly as the algorithm. Only 1 participant was not comfortable with the oncoming vehicle's lateral proximity to the ego and the ego's lane positioning.

SCENARIO 6: A Road with car ahead slowing

Watch the Scenario video, Comfort Algorithm and Scoring here: LINK

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	18	13	12	13	1	17	1
Normal	5	8	12	10	1	7	0
Aggressive	1	5	2	0	0	2	0
Extremely Aggressive	0	0	0	0	0	0	0
N/A	2	0	0	3	24	0	25

Participant overall comfort ratings for Scenario 6: A Road with car ahead slowing

Participant Comfort factor results:

The following No. of participants recorded **Aggressive** comfort ratings:

- Acceleration = 1
- Deceleration = 5
- Headway = 2
- Lane Positioning = 2

There were a 5 (19%) of participants that expressed discomfort at the deceleration of the ego when it approached a slower car ahead, but only 1 participant noted the aggressive acceleration at the beginning of the scenario.

Verbatim analysis:

Here the ego vehicle was driving on an A road, following a car ahead and then slowing down due to a slower car ahead. Participants stated that they felt there was enough distance to the car in front, though many felt that the approach to it was a bit fast and one pressed the trigger as they felt ‘there was a bit of a delay before my car started to slow down’ and ‘approached car in front too quickly’.

Raw data from the comfort algorithm:

Comfort Algorithm Scores: Scenario 6					
	Combined accel	Combined jerk	Lateral Distance (adjacent cars)	Headway	Gap between cars
Seconds	CF1A	CF1B	CF2	CF3	CF4: Not measured
10	19.4	48.4	100	100	100
20	95.45	100	100	100	100
30	91.05	83.05	100	100	100
40	100	100	94	100	100
50	100	100	100	100	100
60	100	100	100	100	100
Averages	84.3	88.6	99.0	100.0	N/A
Comfort Grade	A	A	A*	A*	N/A
Pass/Fail	FAIL	FAIL	PASS	PASS	N/A
Overall comfort score	93.0				
Overall Comfort PASS/FAIL	FAIL	A*			

Scenario 6 Data analysis:

The Comfort algorithm results:

The comfort algorithm reported 2 comfort factor failures:

- Acceleration
- Jerk

These failures are due to aggressive acceleration at the beginning of the scenario.

Overall comfort scoring:

	Participant Results		Algorithm Results		Scoring Matrix		
No. of	**Percentage**	**Overall Comfort Rating**	**Comfort Score**	**Pass/ Fail**	**Algorithm Grades**	**Algorithm Ratings**	**Overall Ave Score**
4	15%	Extremely comfortable			A*	Extremely comfortable	>90
9	35%	Very comfortable			A	Very comfortable	>80
7	27%	Comfortable			B	Comfortable	>75
4	15%	Acceptable			C	Acceptable	>65
2	8%	Uncomfortable			D	Poor	>60
0	0%	Very Uncomfortable		**FAIL**	F	Fail / uncomfortable	<60
26							

Participants:

92% of the participants, rated the scenario experience as Acceptable, Comfortable, Very comfortable or Extremely comfortable.

The highest grouping was 35% rating the scenario as Very Comfortable

2 (8%) of the participants rated the Scenario as, Uncomfortable.

Comfort Algorithm:

The Comfort algorithm generated a FAIL.

Alignment:

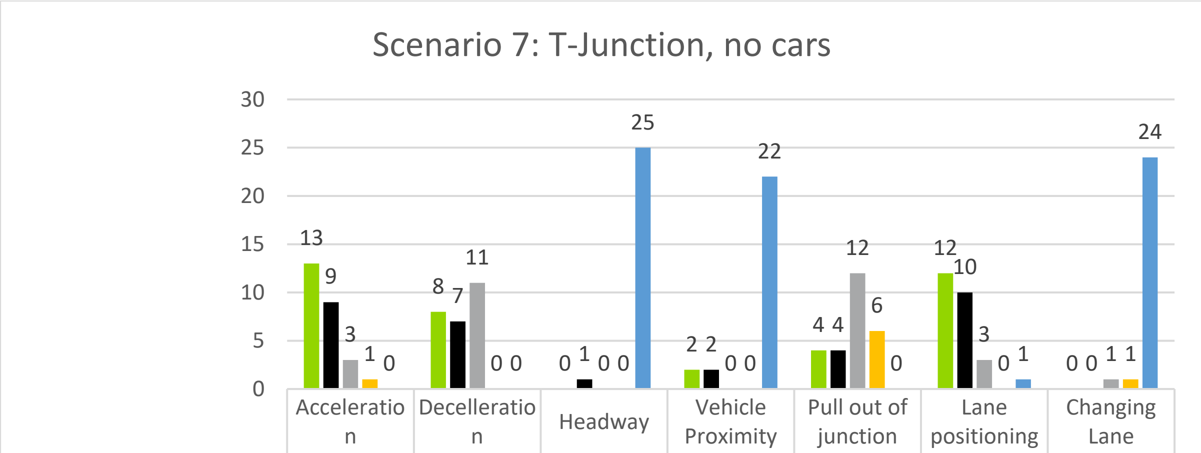
No Alignment. 23 participants said that the acceleration was comfortable or normal. The algorithm failed the ADS in this scenario because of the harsh acceleration at the start of the scenario. This highlights that acceleration is difficult to gauge using the configuration used for the trial. However, it was interesting that the deceleration caused 5 (19%) of the participants to record an 'Aggressive' rating. This is likely due to the participant focusing dead ahead on the road and being able to detect a closing speed that they were not comfortable with. (perhaps if there were more visual cues at the start of the scenario, participants could have detected their acceleration rate better?)

If you exclude the first 4 seconds of the trial, then the results do correlate with 50% of participants rating A or A*.

SCENARIO 7: T-Junction, no cars

Watch the Scenario video, Comfort Algorithm and Scoring here: LINK

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	13	8	0	2	4	12	0
Normal	9	7	1	2	4	10	0
Aggressive	3	11	0	0	12	3	1
Extremely Aggressive	1	0	0	0	6	0	1
N/A	0	0	25	22	0	1	24

Participant overall comfort ratings for Scenario 7: T-Junction, no cars

Participant Comfort factor results:

The following No. of participants recorded **Aggressive** comfort ratings:

- Acceleration = 3
- Deceleration = 11
- Pulling out of junction = 12
- Lane Positioning = 3
- Changing Lane 1

The following No. of participants recorded **Extremely Aggressive** comfort ratings:

- Acceleration = 1
- Pulling out of lane = 6
- Changing lane = 1

Verbatim analysis:

This involved turning right at a T junction with no other cars, however the most frequent feedback from participants was that it was an uncomfortable turn that felt 'not very natural' and 'not human like'. Some participants reported having pulled the trigger. Some stated that they expected the car to stop and that it pulled off while they were checking for oncoming cars at the junction.

Raw data from the comfort algorithm:

Comfort Algorithm Scores: Scenario 7

	Combined accel	Combined jerk	Lateral Distance (a	Headway	Gap between cars
Seconds	CF1A	CF1B	CF2	CF3	CF4: Not measured
10	98.5	100	100	100	100
20	92.25	75.2	100	100	100
30	99.3	99.2	100	100	100
Averages	96.7	91.5	100.0	100.0	N/A
Comfort Grade	A*	A*	A*	A*	N/A
Pass/Fail	PASS	PASS	PASS	PASS	N/A
Overall comfort score	97.0				
Overall Comfort PASS/FAIL	**PASS**	A*			

Scenario 7 Data analysis:

The Comfort algorithm results:

The comfort algorithm reported 0 comfort factor failures

Overall comfort scoring

	Participant Results			Algorithm Results			Scoring Matrix		
No. of	**Percentage**	**Overall Comfort Rating**		**Comfort Score**	**Pass/ Fail**		**Algorithm Grades**	**Algorithm Ratings**	**Overall Ave Score**
1	4%	Extremely comfortable		A*	**PASS**		A*	Extremely comfortable	>90
4	15%	Very comfortable					A	Very comfortable	>80
6	23%	Comfortable					B	Comfortable	>75
11	**42%**	**Acceptable**					C	Acceptable	>65
4	15%	Uncomfortable					D	Poor	>60
0	0%	Very Uncomfortable					F	Fail / uncomfortable	<60
26									

Participants:

85% of the participants, rated the scenario experience as Acceptable, Comfortable, Very comfortable or Extremely comfortable.

4 (15%) of the participants rated the Scenario as, Uncomfortable.

The highest grouping for the overall rating was 42% Acceptable

Comfort Algorithm:

The Comfort algorithm generated a PASS with an A* grade (Extremely comfortable).

Alignment:

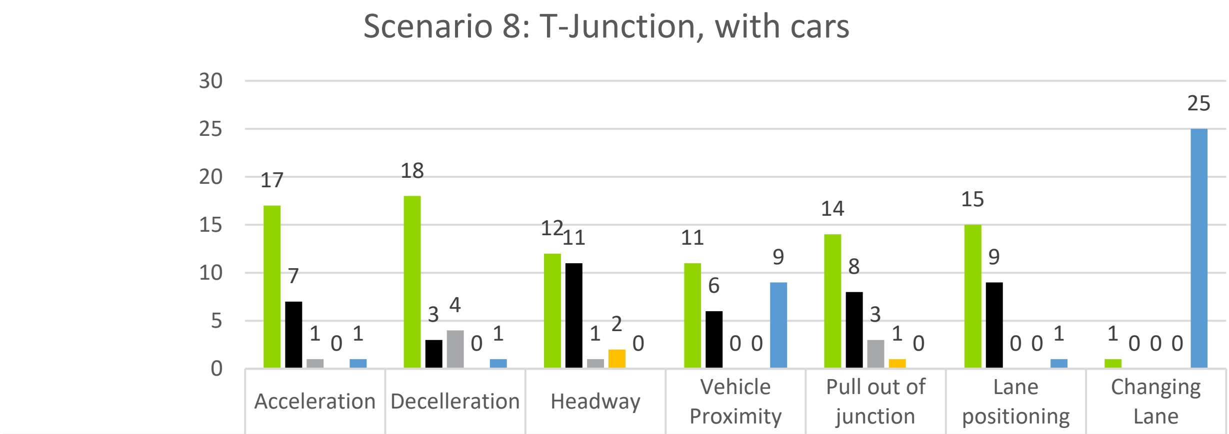
No Alignment. Only 1 participant rated the experience as Extremely comfortable.

The algorithm does not punish sufficiently the harshness of braking and turning angles/rate of directional change.

SCENARIO 8: T Junction, with cars

Watch the Scenario video, Comfort Algorithm and Scoring here: LINK

Participant comfort factor ratings:



	Acceleration	Decelleration	Headway	Vehicle Proximity	Pull out of junction	Lane positioning	Changing Lane
Comfortable	17	18	12	11	14	15	1
Normal	7	3	11	6	8	9	0
Aggressive	1	4	1	0	3	0	0
Extremely Aggressive	0	0	2	0	1	0	0
N/A	1	1	0	9	0	1	25

Participant overall comfort ratings for Scenario 8: T Junction, with cars

Participant Comfort factor results:

The following No. of participants recorded **Aggressive** comfort ratings:

- Acceleration = 1
- Deceleration = 4
- Headway = 1
- Pulled out of junction = 3

The following No. of participants recorded **Extremely Aggressive** comfort ratings:

- Headway = 2
- Pulled out of junction = 1

Verbatim analysis:

Here the vehicle turned right at a T junction with cars approaching the T junction in both directions, and participants overwhelmingly reported that they expected it to stop and it didn't. The deceleration was questioned with comments that it was 'aggressive' and 'hit the brakes quite early'. Some participants stated that they didn't pull the trigger as they 'felt fine'.

Raw data from the comfort algorithm:

	Combined accel	Combined jerk	Lateral Distance (a	Headway	Gap between cars
Seconds	CF1A	CF1B	CF2	CF3	CF4: Measured mai
10	60.75	77.1	100	100	100
20	58.25	97.95	100	100	96
30	68.6	92.2	100	100	100
Averages	62.5	89.1	100.0	100.0	98.7
Comfort Grade	D	A	A*	A*	A*
Pass/Fail	FAIL	PASS	PASS	PASS	PASS
Overall comfort score	90.1				
Overall Comfort PASS/FAIL	FAIL	A*			

Note: 'CF4 - Gap between cars': There was a 10 second gap between the 2 cars approaching from the left. The Ego pulled out 2 secs after the first car passed, leaving a gap of 8 seconds to the next car coming from left. Therefore the 'Normal' rating applied for a gap for 8 seconds between cars. The exposure time from junction to the ego being on the new carriageway was 4 seconds, therefore 4X-1 = -4 points.

Scenario 8 Data analysis:

The Comfort algorithm results:

The comfort algorithm reported 1 comfort factor failure:

- Combined acceleration. (Acceleration and Deceleration)

Overall comfort scoring

	Participant Results			Algorithm Results			Scoring Matrix		
No. of	Percentage	Overall Comfort Rating		Comfort Score	Pass/ Fail		Algorithm Grades	Algorithm Ratings	Overall Ave Score
2	8%	Extremely comfortable					A*	Extremely comfortable	>90
10	**38%**	**Very comfortable**					A	Very comfortable	>80
9	35%	Comfortable					B	Comfortable	>75
1	4%	Acceptable					C	Acceptable	>65
4	15%	Uncomfortable					D	Poor	>60
0	0%	Very Uncomfortable			**FAIL**		F	Fail / uncomfortable	<60
26									

Participants:

85% of the participants, rated the scenario experience as Acceptable, Comfortable, Very comfortable or Extremely comfortable.

4 (15%) of the participants rated the Scenario as, Uncomfortable.

Comfort Algorithm:

The Comfort algorithm generated a FAIL

Alignment:

No Alignment. 46% of participants rated the experience as Very or extremely comfortable. Participants did not highlight any issues with braking or turning which the algorithm implies.

4 people rating the pulling out at junction as Aggressive or extremely aggressive. The numbers/percentages of people who accept a gap of 8-10seconds is in line with the lit. A key consideration is the general traffic levels on the road. People will tend to wait till traffic has passed if there is a larger gap/lack of traffic after a cluster of cars.

Appendix B

Participant demographics

Participants were recruited by opportunity sampling, through sending emails out to colleagues at Connected Places Catapult and also recruiting family and friends.

27 participants completed the trials, 16 of whom were male and 11 of whom were female. The age range was 21 years and above. All but 5 participants have held a driving license for over 21 years (four of whom were 21-30 years old and one of whom was 31-40 years old).

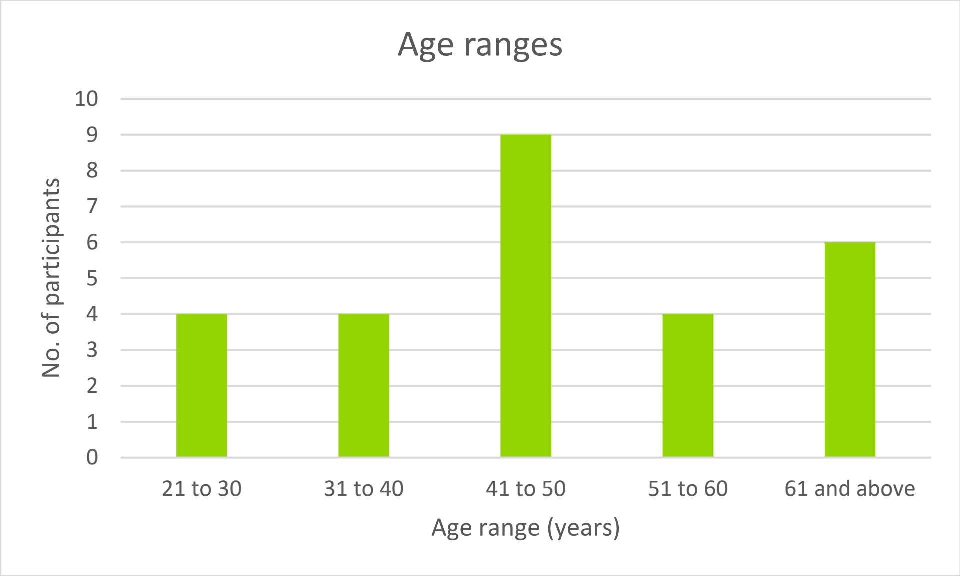


Figure 38: Frequency of participants' age ranges.

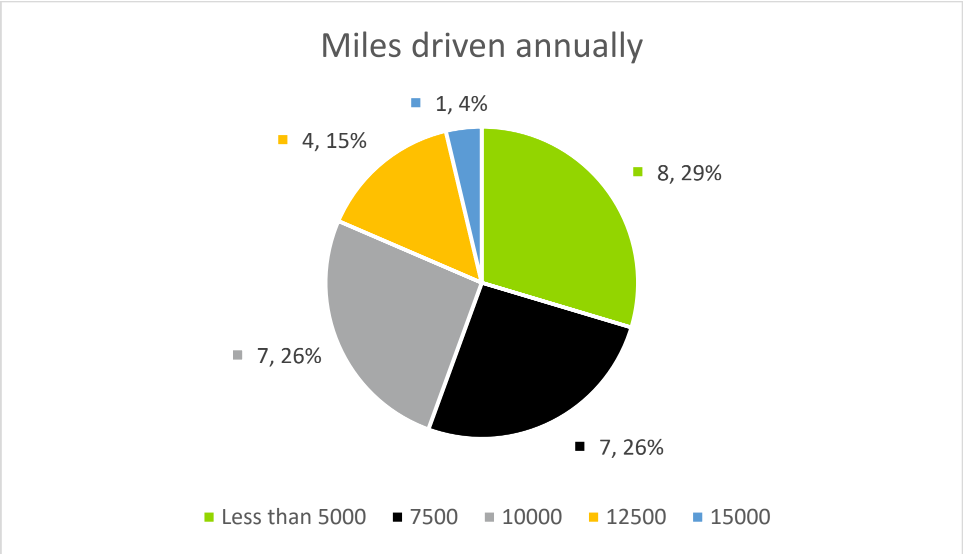


Figure 39: Average annual mileage driven by participants (participant number and percentage)

Thematic Analysis

After each scenario participants were asked to summarise what they had experienced, and their comments were recorded. An inductive method was selected to thematically analyse these comments, i.e. coding and theme development was informed by the content of the data. In order to do this, all comments were collated. The following processes were then used:

1. **Knowledge of the data -** reading and re-reading the data, to become familiar with its content.
2. **Generating initial themes -** examining the collated data to identify significant broader patterns of meaning i.e. themes.
3. **Reviewing themes -** checking the candidate themes against the dataset, to determine that they tell a convincing story of the data.
4. **Coding -** coding the entire dataset by noting which themes occur in each piece of data (Braun and Clarke, 2006).The identified themes were as follows – the figure in brackets after the theme is the number of the theme that was allocated to it for the subsequent statistical analysis:

Physical factors of the vehicle	**Participant behaviour**	**Vehicle behaviour with regards to other cars**	**Vehicle behaviour with regard to the road**
No wing mirrors/couldn't see behind (1)	Pulled trigger (2)	Too close to car in front (4)	Happy with speed (9)
Missing visual indication of signal (13)	Didn't pull trigger (3)	Enough distance to car in front (5)	Not happy with speed – too fast (10)
	Looked at speedometer (8)	Too close to passing cars (6)	Good lane positioning (11)
		Enough distance from passing cars (7)	Poor lane positioning (12)
			Uncomfortable lane change (14)
			Uncomfortable turn (15)
			Expected the vehicle to stop and it didn't (16)

On completing the coding, it was noted that each comment had a maximum of 4 unique themes in its content. There were a small amount of comments (28) for which the content did not pertain to any of the themes, so these were not coded.

The most frequent themes overall are summarised in the below pie chart:

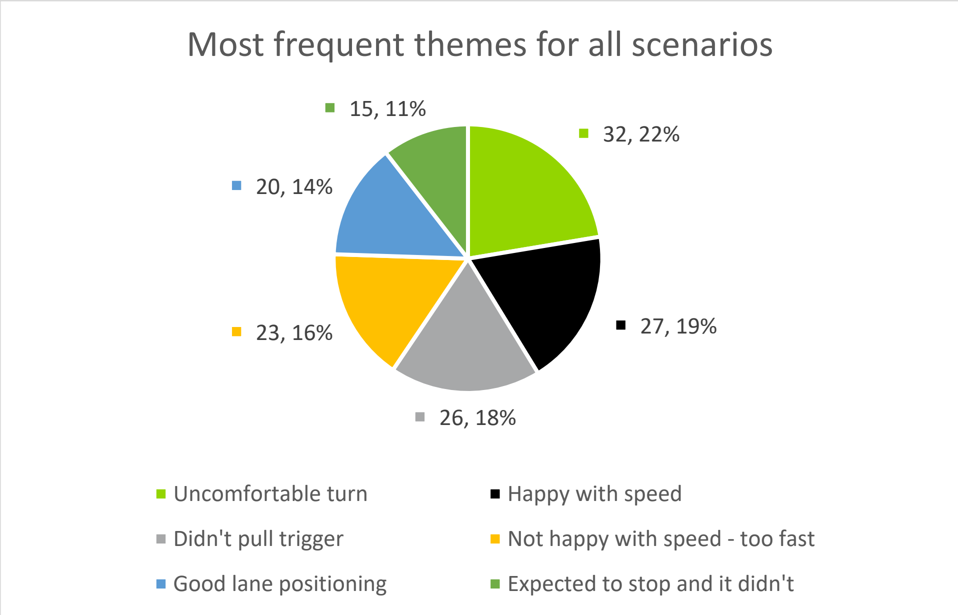


Figure 40: The most frequent themes for all scenarios combined

A summary of the themes per individual scenario is as follows:

Scenario 1:

This scenario involved merging onto a motorway, and the most common theme from qualitative comments were that there were no wing mirrors and as a result participants couldn't see behind them. Some participants voiced that this felt limiting and they tried to move their head more to check the inside lane, as they 'couldn't see cars coming/overtaking till they were next to you'. This lack of situational awareness was cited as the reason that some participants felt uncomfortable, and there was a comment made that 'it used indicators, but wanted to verify myself' and 'having a mirror or screen would help comfort'. There were some comments stating that the ego vehicle 'got close to the car ahead' and some participants reported pulling the trigger as a result.

Scenario 2:

This was changing lanes on a motorway and then taking a motorway exit. There was a mixed spread of themes with no clear prevalent theme. Participants largely summarised what occurred during the scenario, commenting that there was 'no change in acceleration or deceleration as it pulled off' and that they 'felt comfortable at all times, didn't pull the trigger'.

Scenario 3:

Here the AV joined a roundabout and then exited on the second turn-off with no other cars present. Participants' most prevalent comments were that they were uncomfortable with the joining of the roundabout (or turning if they perceived it as a T junction). They perceived the turns as 'very sharp' and ‘very robotic and spinning on the spot’. Some participants also expressed surprise that it didn’t come to a stop at the junction and that they felt like it ‘went almost on the other side of the road’.

Scenario 4:

As with scenario 3 this involved joining a roundabout and exiting, though there were also 2 cars on the roundabout. Participants stated they were happy with the speed that the ego vehicle travelled at, but they expected it to stop and give way to the oncoming vehicle and it didn’t. They stated that ‘I would have waited for the car to pass before I would have pulled out’, and ‘not sure I would have gone ahead’. Some recognised that it was a very similar scenario to the previous one but they felt the turning was ‘smoother’ and that ‘it didn't come out as far into the road as it went round the corner’ (as the order was counterbalanced, this would only apply to those who experienced Scenario 3 before Scenario 4).

Scenario 5:

This scenario involved driving on an A road and passing an oncoming car, after which participants summarised that they were happy with the speed of travel. Participants observed that there was a car coming the other way but stated that overall, the scenario was ‘uneventful’ and a ‘nice comfortable easy drive’. A small number of participants found the second turning ‘sharp’ and one pulled the trigger as a result of this.

Scenario 6:

Here the ego vehicle was driving on an A road, following a car ahead and then slowing down due to a slower car ahead. Participants stated that they felt there was enough distance to the car in front, though many felt that the approach to it was a bit fast and one pressed the trigger as they felt ‘there was a bit of a delay before my car started to slow down’ and ‘approached car in front too quickly’.

Scenario 7:

This involved turning right at a T junction with no other cars, however the most frequent feedback from participants was that it was an uncomfortable turn that felt ‘not very natural’ and ‘not human like’. Some participants reported having pulled the trigger. Some stated that they expected the car to stop and that it pulled off while they were checking for oncoming cars at the junction.

Scenario 8:

Here the vehicle turned right at a T junction with cars approaching the T junction in both directions, and participants overwhelmingly reported that they expected it to stop and it didn’t. The deceleration was questioned with comments that it was ‘aggressive’ and ‘hit the brakes quite early’. Some participants stated that they didn’t pull the trigger as they ‘felt fine’.

Inter-item Correlation scores

Correlation coefficient formulae were used to find how strong the relationship was between items of a question. The formulae returned a value between -1 and 1. 1 indicates a strong positive relationship, -1 indicates a strong negative relationship and a result of zero indicates no relationship at all.

Inter-Item Correlation Matrix

	ATTAUTO_01	ATTAUTO_02	ATTAUTO_03	ATTAUTO_04	ATTAUTO_05	ATTAUTO_06	ATTAUTO_07	ATTAUTO_08
ATTAUTO_01	1.000	.389	.302	-.324	.176	-.181	.007	-.019
ATTAUTO_02	.389	1.000	.600	-.487	-.230	-.171	-.165	-.195
ATTAUTO_03	.302	.600	1.000	.013	.004	.236	-.218	-.294
ATTAUTO_04	-.324	-.487	.013	1.000	.162	.293	-.138	-.228
ATTAUTO_05	.176	-.230	.004	.162	1.000	.688	.762	.602
ATTAUTO_06	-.181	-.171	.236	.293	.688	1.000	.527	.363
ATTAUTO_07	.007	-.165	-.218	-.138	.762	.527	1.000	.712
ATTAUTO_08	-.019	-.195	-.294	-.228	.602	.363	.712	1.000

Figure 41: Inter-Item Correlation scores – Attitude to Automation

Some questions within each item correlated highly with each other. For example in terms of attitude to automation, item 2 ('Automation helps me to save time') and item 3 ('Automation helps me to be more flexible in terms of time') correlated highly, potentially because they are both related to time. In addition, item 5 ('I mostly feel uncomfortable using a machine') and item 6 ('I have no confidence in using machines') are highly correlated, as the content in both relates to how they feel using machines. Item 7 ('I often have problems with the handling of machines' and item 8 ('Machines often do not do what I want') are correlated highly, potentially as they both relate to ability to work machines.

Inter-Item Correlation Matrix

	ATTCAV_02	ATTCAV_03	ATTCAV_04	ATTCAV_05	ATTCAV_07	ATTCAV_01_Reverse
ATTCAV_02	1.000	.809	-.024	.738	.691	.851
ATTCAV_03	.809	1.000	-.098	.594	.530	.809
ATTCAV_04	-.024	-.098	1.000	.391	.599	.089
ATTCAV_05	.738	.594	.391	1.000	.911	.846
ATTCAV_07	.691	.530	.599	.911	1.000	.801
ATTCAV_01_Reverse	.851	.809	.089	.846	.801	1.000

Figure 42: Inter-Item Correlation scores – Attitude to CAVs

A Cronbachs Alpha test was run as a measure of internal consistency i.e. how closely related a set of items are as a group. It is considered to be a measure of scale reliability. As a result of the test, Item 6 was removed in order to increase reliability, and item 1 was reverse coded to align with the others i.e. so the pattern of positive to negative sentiment was consistent throughout. For Attitude to CAVs, Item 2 (Do you think you would find an automated driving system unreliable or reliable) was highly correlated with Item 3 (Do you think you would find an automated driving system not useful or useful). Item 2 also correlated highly with Item 5 (do you think you would find an automated driving system unsafe or safe), Item 7 (Do you think you would find an automated driving system stressed or relaxed) and Item 1 (Do you think you would find an automated driving system untrustworthy or trustworthy). Item 1 and Item 3 correlated highly, as did 2 and 5, 2 and 7.

There was a significant effect of gender on Attitude to Automation and Attitude to CAVs ($p<0.05$), in which females scored higher on Attitude to Automation and males scored higher on Attitude to CAVs.