A Pilot Study on Situational Trust and Driver's Take Over Performance in High-Level of Autonomous Driving.

Sheng-en, Peng

Department of Cognition, Data and Education

BSc Psychology, University of Twente

First supervisor: Dr. Simone Borsci

Second supervisor: Dr. Rob H.J van der Lubbe.

Date: July 4th, 2023

APA 7th edition

Abstract

In this pilot study on autonomous driving, we explored the importance of situational trust and its effects on various factors such as sleepiness, state anxiety, mental workload, task difficulty, and driver's takeover performance in high-level automated driving. We used the SSAM model and involved six participants who experienced seven scenarios with different levels of difficulty using a driving simulator and VR headset. Participants responded to Takeover Requests (TORs), and we measured situational trust, sleepiness, state anxiety, mental workload, and takeover performance during the handover task. The results indicated that situational trust did not act as a mediator in the proposed SSAM model, suggesting its limited impact on the connections between variables. However, task difficulty was found to play a crucial role in handover tasks.

Acknowledgement

First and foremost, I want to express my deepest appreciation to my research supervisor Dr Simone Borsci, for being my motivator and mentor throughout the thesis project. His dynamism, vision and sincerity have deeply inspired me and added colours to my interest in the field of Human factors and engineering. I would also like to thank him for his great sense of humour, where meetings are never boring. Moreover, I would like to express my sincere gratitude to my second supervisor Dr. Rob H.J van der Lubbe, for his valuable feedback while revising my thesis. His critical insights and constructive comments have greatly contributed to improving the quality of my work. I would not have been able to complete this thesis with full confidence without the support from my supervisors. It was a great privilege and honour to work and study under their guidance.

I would also like to express my heartfelt appreciation and acknowledgement for my amazing teammates, Kerem Doğan and Bart Haagsma. Together, we have formed an exceptional team that has achieved and completed this project. Every one of you has played an integral role in our collective success and I am grateful to have the opportunity to work with you both. What a wonderful journey it has been.

Furthermore, no words cannot express my gratitude to my Enschede guardian angel Karl Hinrichs and Bianca Hinrichs. You have been a constant source of love, guidance, and protection, standing by my side with unwavering support and care. Being an international student during the pandemic, your selfless acts of kindness have provided me with a sense of security and stability that is beyond measure. Your presence in my life has been a beacon of light, illuminating my path and ensuring that I never feel alone. Of course, I will not forget the lovely dogs of theirs Saar, Juul and Noor, who have provided me with unconditional love and been the best study buddy.

Lastly, with heartfelt thanks and immense appreciation, I would like to acknowledge my parents. Who has never doubted me, instead they fully supported me in this academic journey in the Netherlands. Their unconditional love, unwavering support and encouragement have moulded me into the person I am today. Their support in my life is immeasurable. Dear Mama and Papa, your roles as my parents are cherished and treasured deeply.

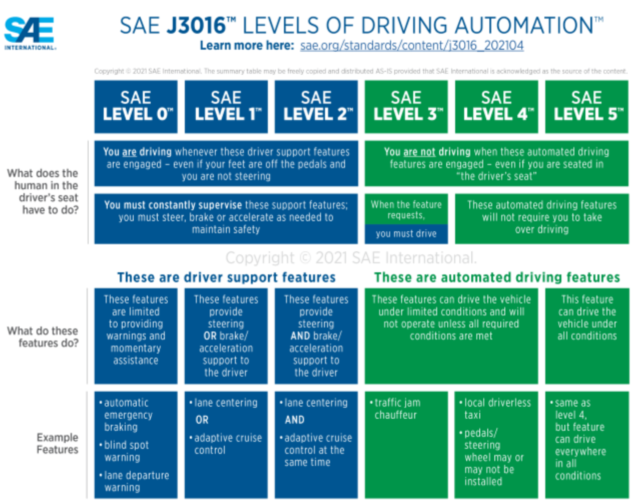
Content

Introduction

With the current automotive technology, for the first time, human drivers can engage with non-driving related tasks (e.g., reading a book) while driving. According to the Society of Automotive Engineers (SAE), these vehicles are classified as "Level 3" or "conditional automation" vehicles (SAE, 2021). Figure 1 presented the overview of all levels of SAE automation. In Level 3 automation, drivers can be temporary "out of the loop", in which the autonomous vehicles can operate independently without interference from human drivers in controlling and monitoring them (Merat et al., 2019). However, drivers are still expected to interfere when the automation reaches its limit (Soares et al., 2021). Driver assistance tools such as Takeover Request (TOR) allow the driver to assume control over the vehicle from the automation system when necessary. A takeover performance gives drivers the opportunity to regain control of the automated system when needed. In contrast, when the automated system transfers control to the driver via TOR, it is known as a handover task.

Figure 1

SAE Levels of driving automation



Note: This figure is derived from (SAE, 2021) on 7th March 2023

Conditional automation benefits will not be realised until Level 3 vehicles are widely accepted and used on the roads. If drivers do not trust the automated system, it is less likely for them to successfully take over in an emergency situation. Failing to react to unforeseen situations that demand immediate takeover would be a concerning scenario, particularly in the context of Level 3 automation where human drivers are less active (Soares et al., 2021) In line with these considerations, trust in automation is seemingly a complex phenomenon affecting drivers' takeover performance. As a psychological construct (Muir, 1994), trust is difficult to manipulate and varies in time (Jin et al., 2021). With the existing literature, there is more than one definition that can be found to explain trust in automation. However, in this thesis, we will adopt the definition of trust in automation by Körber et al. (2018), as this definition explains trust as a multi-dimensional concept, that is built upon the relevant characteristics of the automated system (e.g, predictability of the automation system) and the trustor (e.g, willingness to trust the automation). The definition of trust in automation refers to the willingness of a user to be vulnerable to the actions of automation, irrespective of the ability to supervise or intervene in those actions" (Körber et al., 2018, p. 19).

As proposed by Hoff and Bashir (2015), trust in automation can be divided into three layers: dispositional trust, situational trust and learned trust. Firstly, dispositional trust relates to the operator's characteristics such as age, gender, personality, and culture (Hoff and Bashir (2015). Next, situational trust is determined by the environment under specific situations and timeframes. Lastly, in the layer of learned trust, experience and interaction of automation come into play. Previous studies have primarily focused on situational trust with limited independent variables, often conducted in controlled experimental settings. For instance, Ferraro and Mouloua (2022), discussed the influence of the self-efficacy of drivers towards their situational trust in automation systems. A study by Avetisian et al. (2022), has proven that situational trust leads to a high level of positive emotions in drivers towards automated vehicles. Additionally, age was found to be a factor that impacted the level of situational trust in autonomous driving. Older drivers were found to have a lower level of situational trust and require a higher level of mental effort in autonomous driving (Zheng et al., 2023). However, using controlled experiments alone to study situational trust in autonomous driving is insufficient to fully comprehend the complexity of this concept and its effect on drivers' takeovers performance.

With such motivation in mind, conducting further research to investigate the interaction between multiple factors, situational trust and takeover performance becomes essential. Therefore, this thesis proposed a multidimensional model for exploring the relationship between the level of handover task, takeover performances of human drivers and situational trust in the context of L**evel 3** automation. Moreover, this work aims to explore the effects of sleepiness, state anxiety and mental workload as the sub-factors in the process of automated cars to drivers. The potential contribution of situational trust in the takeover performance and the relationship between sleepiness, state anxiety and mental workloa**d** will be modelled by the Structural Equation Modelling approach (SEM).

Main aspects that can affect handover in automation

Situational Trust and Takeover Performance

Situational trust in automation uncovers the factors influencing drivers' behaviours, decision-making, and performance in specific driving situations, and holds relevance in this study (Hoff & Bashir, 2015). In other words, through studying situational trust it provided insights into how drivers' trust in automation is influenced by specific situations. To examine this, our driving simulator study aims to explore the effects of situational trust in automation on takeover performance, rather than trust in automation in general. As this approach allows us to understand the impact of trust within the context of specific driving situations and subsequent handover processes.

To comprehensively assess sub-factors (sleepiness, state anxiety and mental workload) and situational trust, subjective measurement methods were employed. Since trust is considered an individual’s attitude, questionnaires have been suggested as an effective means to assess trust in automation (Körber, 2018). This approach aligns with the perspective put forth by Mayer et al. (1995), who recommend the use of self-report or similar methods for measuring trust in automation. In addition, subjective measurement methods were also recommended for assessing the sub-factors, as suggested by previous literature. Muckler and Seven (1992) emphasise the importance of subjective measures in evaluating mental workload. Similarly, in a related study on automation vehicles, Chen et al. (2019) employed a subjective measurement technique to assess the mental workload experienced during Level 2 autonomous driving.

Moreover, the study conducted by Cai et al. (2023) elucidates the significance of evaluating subjective sleepiness within the context of driving conditions. The researchers assert that by examining drivers’ self-reported experience of sleepiness, valuable insights can be gained regarding their levels of alertness during the act of driving. The use of subjective measurement on sleepiness can serve as a cue for drivers to identify their fatigue and draw a line when necessary, such as pulling over when needed (Cai et al., 2023).

The subjective measurement has not been explicitly stated as the most recommended way to assess state anxiety in any definitive paper to date. It is noteworthy, however, according to the study of Kim et al. (2022), who investigated the takeover performance of drivers when driving in Level 3 automation, used subjective measures on state anxiety. Thus, subjective measures were deemed appropriate for capturing the drivers' subjective experiences of anxiety in that study. Therefore, a subjective questionnaire was used in this study as a means of capturing situational trust fluctuations and sub-factors throughout the experiment, so that an extensive assessment of the relationship between trust and driver behaviour could be made.

Situational Trust, Sleepiness and Takeover Performance

Previous studies involving taking over control from autonomous cars did not examine the relationship between sleepiness and situational trust. However, previous research suggests that drivers who trust automation more feel fatigued while driving automated vehicles (Kundinger et al, 2019). The results indicate that drivers' level of sleepiness during automated driving is directly influenced by their trust in automation. So, we hypothesised that situational trust might affect sleepiness in autonomous driving, and we investigated this hypothesis in our experiment.

Situational Trust, State Anxiety and Takeover Performance

To dive even deeper into the theoretical foundation of situational trust in automation, state anxiety and takeover performance it would be worth first explaining the concept of state anxiety. Anxiety can be separated into two components: trait anxiety and state anxiety (Spielberger & Smith, 1966). Trait anxiety refers to the inherent characteristics of an individual. Where, state anxiety is defined as a temporary physiological reaction, feelings and pressure associated with an event or a state. In this study, our focus is on state anxiety as we aim to investigate the conditioned pressure that participants experience during the handover task.

Based on a thorough literature search, no evidence supporting the impact of situational trust on state anxiety and takeover performance was found. It has been shown, however, that trust in automation can have a significant impact on state anxiety. According to the study of Lu et al. (2022), the effects of state anxiety on trust, situational awareness, and role adaptation were found to be negatively predictive. In other words, when the driver is experiencing a lower level of state anxiety, trust in automated driving tends to be higher. Similarly, situational awareness and role adaptation will also increase when state anxiety falls, however, these two factors were not the primary topic of this discussion, therefore we will not delve further. Besides, the correlation between trust and state anxiety (and the other two factors), an indirect impact was also found between these two variables. However, there is still information missing from the existing literature. The relationship between state anxiety and trust was only studied with situation awareness and role adaption, other potential factors that could act as a possible influencer are yet unknown. Recognising the existing gap in knowledge, it is essential to investigate the potential influence of situational trust on state anxiety and takeover performance. Therefore, we posit there may be a relationship between situational trust, state anxiety, and takeover performances in autonomous driving scenarios.

Situational Trust, Mental Workload and Takeover Performance

As with the other two factors, there is no evidence that situational trust directly impacts mental workload and handover performance. However, the study by Clement et al. (2022) has proven the important relationship between trust and mental workload. Their experimental result has suggested that trust negatively affects a driver’s mental workload while driving. To further explain, drivers with greater levels of trust in automation tend to have lower mental workloads when driving. Similarly, Du et al. (2019), have also found that, when the driver’s trust is initially higher, the mental workload has resulted to be lower. Moreover, Yousfi et al. (2021) found that trust in automation can be affected by the time given to drivers to take over, as well as trust has a negative predictive effect on mental workload. Although mental workload has been extensively studied, it is still unclear what effect situational trust had on mental workload and handover performance. Therefore, we hypothesised that situational trust would have a significant influence on mental workload and takeover performance. Considering this hypothesis, we undertook a comprehensive investigation in our experiment to explore the relationship between situational trust, mental workload, and takeover performance.

Situational Trust, Sub-Factors and Handover Task

The interplay between situational trust, sub-factors (sleepiness, state anxiety and mental workload) and the handover tasks in autonomous driving scenarios remains unexplored in the existing literature. While studies have separately examined these factors, their interconnectedness and influence on each other have not been thoroughly investigated. To bridge this knowledge gap, our study takes a comprehensive approach by including an exploratory question that seeks to unravel the relationship between situational trust, subfactors and the level of difficulty in handover tasks. We argue that the level of difficulty in the handover task will affect situational trust and the sub-factors. Therefore, we aim to enhance our understanding of how the level of difficulty in handover tasks affects situational trust and sub-factors in the context of autonomous driving.

Purpose of the present study

Our goal in this pilot study is to evaluate the effects of varying levels of difficulty in the handover task from the automation system to the driver. Within the context of Level 3 autonomous vehicles, we aim to examine how the level of difficulty impacts situational trust and takeover performance. Building upon previous literature, our study assumed (as summarised in Figure 2) that success in taking over is influenced by situational trust (Figure 2, b) and that situational trust would be influenced by the level of difficulty of the handover task (Figure 2, a). In addition to that, we also expect the level of difficulty will affect success in taking over (Figure 2, c). These assumptions are summarised in the Model 1 (Figure 2).

Figure 2

Conceptual Framework of Model 1

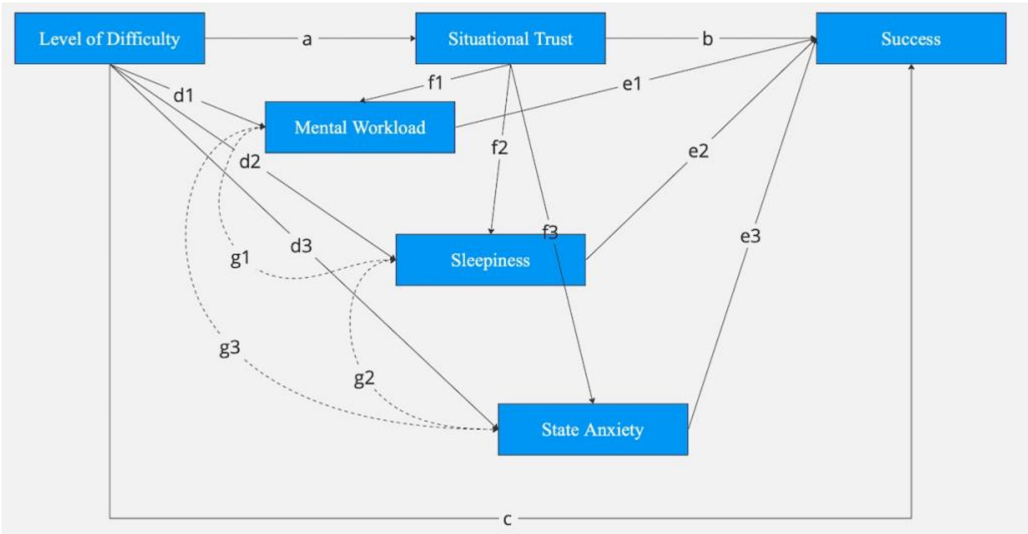
Note. In this figure, "Success" refers to the driver's performance in taking over.

Moreover, drawing from the existing literature (Kundinger et al., 2019; Lu et al., 2022; Clement et al., 2022; Yousif et al., 2022; Du et al, 2019), the sub-factors (Sleepiness, State Anxiety, and Mental workload) factors have been acknowledged to be linked with takeover performances in the context of Level 3 automation driving. A more comprehensive approach is warranted, since studying situational trust alone may not fully capture the intricacies of the relationship in the context of Level 3 automation driving. Recognising the multifaceted nature of driver-automation interaction, it is essential to explore additional factors contributing to this relationship's complexity. By considering a broad range of factors, our study aims to examine the effects of situational trust on the sub-factors. It also examines the effects of the sub-factors on takeover performance. We, therefore formulated the SSAM model, illustrated in Figure 3, by incorporating the sub-factors into Model 1 presented in Figure 2. By applying the SSAM model, the present study aims to investigate the following research question:

RQ: *Does the driver's subjective tendencies towards situational trust, sleepiness, state anxiety, and mental workload influence their takeover performance during handover tasks in the context of SAE-L3 autonomous driving?*

Figure 3

SSAM Model



The graphical representation in Figure 3 summarised the following expectation:

H1: The driver's situational trust will be negatively impacted by the level of difficulty in handover situations. (Figure 3, a) This is in line with the indication coming from Yousfi et al (2022).

H2: The driver's performance in take-over will be positively impacted by situational trust. (Figure 3, b)

H3: The takeover performance will be negatively affected by the level of difficulty (Figure 3, c)

H4: In expectation, the mental workload will be negatively impacted by the level of difficulty in the handover task. (Figure 3, d1)

H5: In expectation, sleepiness will be negatively impacted by the level of difficulty in the handover tasks. (Figure 3, d2)

H6: In expectation, state anxiety will be negatively impacted by the level of difficulty in handover tasks. (Figure 3, d3)

H7: In expectation, the driver's takeover performance will be negatively impacted by the mental workload (Figure 3, e1)

H8: In expectation, the driver's takeover performance will be negatively impacted by sleepiness. (Figure 3, e2)

H9: In expectation, the driver's takeover performance will be negatively impacted by state anxiety. (Figure 3, e3)

H10: In expectation, the mental workload will be negatively impacted by the situational trust (Figure 3, f1). This is aligned with the study of Clement et al. (2022) and Yousfi et al. (2022)

H11: In expectation, sleepiness will be positively impacted by situational trust (Figure 3, f2). This is derived from the study of Kundinger et al. (2019).

H12: In expectation, state anxiety will be negatively impacted by situational trust (Figure 3, f3). This is in line with the findings of Lu et al. (2022)

H13: In expectation, the mental workload is negatively related to sleepiness (Figure 3, g1)

H14: In expectation, sleepiness is negatively related to state anxiety (Figure 3, g2)

H15: In expectation, the mental workload is positively related to state anxiety (Figure 3, g3)

In this pilot study our aim is to assess whether the model presented in Figure 3 (SSAM Model), which has been derived from existing literature, adequately fits the collected data, and can effectively address the research question. By conducting this analysis with a small sample of participants, we can gain insight that will inform and guide future research.

Method

Participants

A total of six participants were recruited for the study through voluntary sampling from the University of Twente. The sample consisted of three males and three females aged between 20 and 25 (Mean = 21.6, SD = 1.94). All participants were licensed drivers with an average of 3.34 years of driving experience (SD = 1.63). There are several requirements involved in the selection of participants in this experiment. Firstly, to participate in this experiment, all participants needed to have a valid driving licence and driving experience. Secondly, participants who have consumed alcohol within the past 24 hours were excluded from this study. The consumption of alcohol was found to be hindering the performance of subjects during the experiment setting. Similarly, participants who consumed caffeine five hours before the experiment were withdrawn, as this study intended to measure the sleepiness of the participants during the experiment, and caffeine consumption affects the data collection on sleepiness. Moreover, participants who have experience with motion sickness were excluded, since this study conflated the use of a driving simulator and VR Head headsets, therefore we expected that participants might experience a certain level of motion sickness during the experiment. In addition, female participants who are pregnant or at risk of pregnancy were excluded from this study, as the use of driving simulators and VR headsets could bring potential risks to the foetus and expectant mothers. Participants with colour blindness and other eye conditions (e.g., Presbyopia, Amblyopia) that might affect the interaction with the virtual system were withdrawn from the study as well. No participant withdrew from this study as all recruited participants met the experiment criteria.

Design

This experiment followed a with-in-subject design with as a dependent variable the success of participants in taking over the command from the car and in executing the task and as the independent variable the level of difficulty of Handover tasks. The sub-variables such as situational trust, sleepiness, mental workload and state anxiety were considered as the mediators.

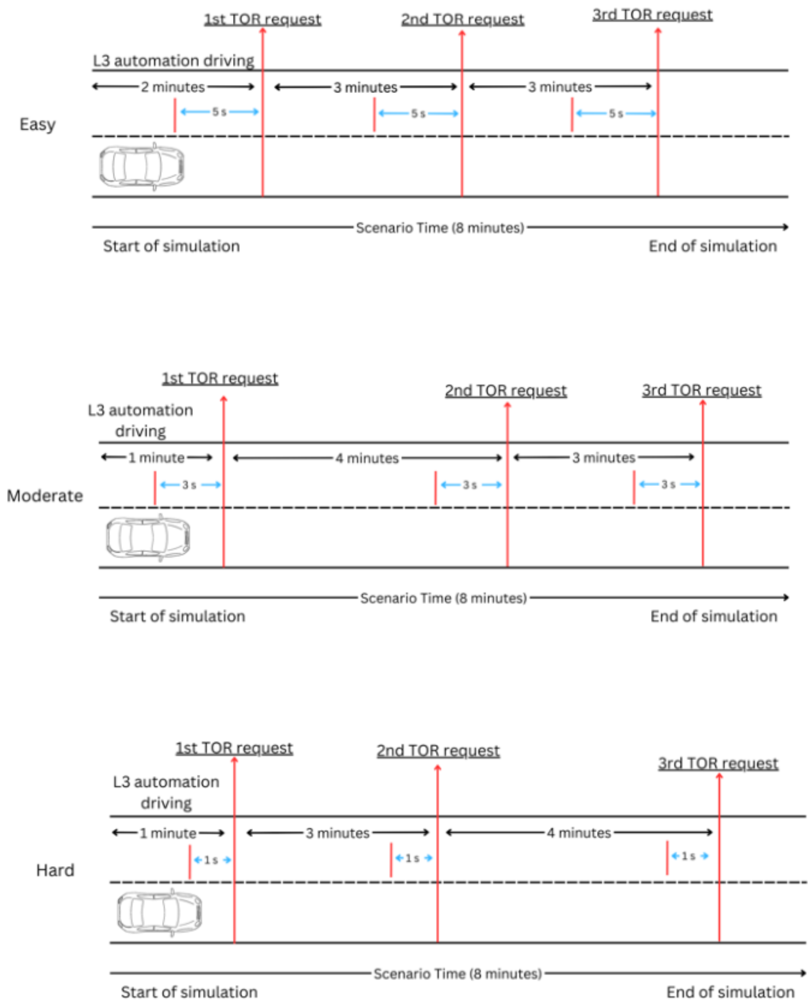
Materials

Instruction and Scenarios

This study involved participants performing a simulated driving task where they had to respond to the TOR displayed through the VR headset. A total of seven scenarios were presented to the participants, and they were asked to follow the TOR instructions such as steering left or right or breaking. The task consisted of three levels of difficulty (Easy, Moderate, and Hard) with a training phase and a test phase. Participants had a different amount of time to react to TOR instructions depending on the conditions, with easier conditions allowing more time and harder conditions providing less. Each scenario lasted eight minutes, with participants receiving instant feedback after each TOR task. In the feedback, they were informed if they had completed the requested action successfully or not. Detailed information on the different conditions and their corresponding actions can be found in Appendix B Table B1. Visualisation of each condition is presented in Figure 4.

Figure 4

Level of difficulty timeline



Note. The "s" indicates second. The blue arrows indicated the window of opportunity for participants to react.

Driving Simulator

This study utilized a high-fidelity advanced driving simulator (Cubicus B200VK) from the University of Twente. The simulator provided a realistic driving experience for the participants. It includes a steering wheel and pedal (Logitech G920 Driving Force) where the participant takes control of the simulated vehicles. A "next-level racing chair" was built into the driving simulator, in which the participant can adjust the seating position based on their comfort preference. Participants were expected to actively engage with the pedals and steering wheel throughout the experiment, as they were expected to react accordingly to the given instruction. Observation is used to record the performance of participants, regardless of whether they achieved success in the TOR task. The automated mode of the driving simulator system was programmed by Unity, where the simulated vehicles perform a Level 3 function during the experiment. See Appendix A, Figures A1, A2, A3, and A5 for an overview of the experiment setup.

Virtual Reality Headset

A Virtual reality headset, Varjo XR-3 were used in this study to provide the participant with a realistic artificial environment while driving the simulated vehicles (see Appendix A, Figure A4). The VR headset will work accordingly with the driving simulator, in other words, the VR headset will change the view of the driving environment based on the movement of the simulated vehicle. See Appendix A, Figure A6 for the VR visualisation.

Experimenter Manual

To ensure the experiment is conducted as organised and with standard, an experimenter manual was created. The manual contains a verbal introduction script where the experimenter can verbally explain the outline of the experiment including, the research aims, duration of the experiment, and several points that the participant need to be aware of before, during and after the experiment. For instance, participants were required to place their phones away from their pockets in silence mode, as they might be distracted by it. One important point of using the manual is to remind the experimenter to always inform the participant that they might experience motion sickness due to the use of a VR headset and driving simulator. For an overview of the manual see Appendix C.

Observation Sheet

To collect data on the participant's performance during the handover task, an observation sheet in Excel was used. If the participant successfully takes over the task, the experimenter will input "1" in the sheets, and "0" for failure to take over the task. An overview of the observation sheet is included in Appendix D.

Subjective Measurement Questionnaire

A subjective measurement questionnaire was used in this study to collect participants' sleepiness, state anxiety and mental workload towards the TOR during the experiment. To reduce the use of paper in this experiment, the questionnaire was created on Qualtrics (an online survey system) and presented to the participant on a tablet (Microsoft Surface Pro 9) (View Appendix E, Figure E1). Informed consent was attached at the beginning of the questionnaire, which was used as a legal agreement of the participants that they are well-informed and fully understand the aim and purpose of the study. It also demonstrated that the participants understood their rights in the study, and they can withdraw anytime without providing any reasons. Moreover, the researchers' and supervisors' contact information can be found on the consent, where participants can contact the researchers if they have any questions in future. The participants filled in their demographic information (e.g., age, sex, and gender) and indicate whether they fulfil the experiment requirement. At this point, only female participants received the questionnaire with the option to indicate their pregnancy status, whereas male participants received the questionnaire without such an option.

To measure the participant's anxiety level before the experiment (also known as trait anxiety), a four-item questionnaire from Spielberger et al. (1971) (as cited in Lu et al., 2022) was derived, which the participants can indicate their level of anxiety from Strongly disagree (1) to Strongly agree (7). Trait anxiety was compared with the state anxiety of the participant in data analysis. Then, a six-item Cyber Sickness in Virtual Reality Questionnaire (CSQ-VR) (Kourtesis et al., 2023) was used to determine whether the participants are experiencing symptoms of motion sickness from the use of a VR headset during the experience from Absent feeling (1) to Extreme feeling (7).

To measure the sub-factors, several subjective measurements were used. Firstly, to measure the subjective level of sleepiness of the participants during the experiment we obtained the nine-point scale measurement, Karolinska Sleepiness Scale (KSS) by Åkerstedt and Gillberg (1990). Participants can indicate their sleepiness in the range from Extremely alert (1), to Very sleepy, great effort to keep awake, fighting sleep (9). Secondly, to measure participants' situational trust level towards autonomous vehicles a one-item Likert scale was used, ranging from Strongly disagree (1) to Strongly agree (7). We adapted this situational trust measurement from Lu et al. (2022). Thirdly, the state anxiety towards the TOR was measured by a three-item scale (Lu et al., 2022) with a range of Strongly Disagree (1) to Strongly Agree (7). Lastly, the mental workload during the whole TOR task will be measured using the Rating Scale Mental Effort (RSME), that are attached in Appendix B, Figure B1. The RSME scale has been described by Zijlstra (1993) as a method of measuring mental effort spent on a task by an individual. Participants had to indicate how much mental effort they had placed on completing the handover process task on a scale of 0 to 150 marked within nine anchor points. The summary of the survey is enclosed in Appendix B, Table B2 the evaluation criteria are presented.

Procedure

Participants were greeted warmly upon arrival and given spoken instructions that contained the overview of the experiment. In this introduction, participants were informed of their rights, which they can withdraw at any time without explanation. Moreover, we made clear to the participants that they might feel uncomfortable or experience motion sickness due to the use of a VR headset and a driving simulator. Due to this, participants could pause the experiment at any moment, if they experienced any symptoms of motion sickness (e.g., dizziness, feeling to vomit). Once making sure the participants have no questions, they were directed to the driving simulator, more instructions on how to respond to the TOR were provided (e.g., steer to the left or right, Middle paddle for the break) and several points for participants to keep in mind during the experiment, for more detail please refer to Appendix C.

Next, the participants were given a tablet on which the informed consent was presented, and they can respond to the questionnaire by clicking the boxes. The demographic information (Participant Number, Age, Sex, and Gender) was filled in by the participants and indicated whether they fit the experiment criteria. At this stage, only female participants received the criteria questionnaire, where they indicated their pregnancy status. On the other hand, male participants received a criteria questionnaire without such an option.

The entire experiment ran in parallel with the use of the questionnaire. Before the experiment began, it was required that the participants filled in their trait anxiety and CSQ-VR to make sure that they were not in a state of motion sickness. Next, the questionnaire instructed the participant to return the tablet and prepare for the scenarios. After participants verbally confirmed their readiness, the experimenter placed the VR headset on for the participants. When the VR headset is adjusted to the best position, experimenters began to calibrate on iMotion, by doing so the eye movement of participants during the experiment can be captured and recorded. After calibration, the participants began with the TOR task with Unity. Before starting the task, the experimenter made sure the participants' hands were not on the steering wheel, as this is identical to the setting of SAE-L3 automation. The participant's dominant feet were instructed to place on the middle paddle this is not identical to the setting of SAE-L3, however, due to vision restrictions from VR headsets, it can be difficult for the participant to brake during the task. Thus, participants were instructed to do so.

This experiment contained 7 scenarios in total, which were divided between the test phrase (1 scenario, 2 TOR tasks) and the test phase (6 scenarios, 3 TOR tasks in each scenario). The participants were observed by the experimenter to identify whether they successfully take over or failed to take over. If the participants performed the required action (e.g., Change Lane to left) within the window of opportunity (e.g., 3 seconds) it will be considered a success, and they will be verbally noticed. Similarly, if the participant did not perform the required tasks within the window of opportunity, it failed. Experimenters used thumbs up to indicate a successful takeover and thumbs down to indicate a failed takeover to each other.

After each scenario, the VR headset was removed, and participants filled in the questionnaire. Then, the experimenter verbally confirmed with the participant that they are not suffering from motion sickness. The experiment continued if the participant does not experience motion sickness. This process will be repeated till the last scenario (7th scenario). If the participants were experiencing motion sickness, they did not continue with the experiment and further assistance was being provided. The entire experiment took about 90 minutes to complete. After three or four scenarios, the experimenter suggested a 5-minute break and offer the participant water for refreshment. During the break, participants were encouraged to walk around in the experiment room, the door of the simulator room was open for ventilation, these were done to avoid motion sickness. Finally, once the participants completed all tasks, they are warmly thanked for their participation and informed again with the contact details in case they have questions or wish to withdraw from the experiment. An illustration of the process can be found in Appendix F, Figure F1.

Data analysis

In this study, R studio (version 2023.06.0+421) was used to analyse the collected data. Moreover, to control the motion sickness a linear regression was performed to examine the relationship between motion sickness and the level of difficulty, and finally, we also controlled by a regression whether trait anxiety of people predict levels of state anxiety. Additionally, descriptive statistics were calculated among the variables and level of difficulty.

A manipulation check was run to explore the influence of the level of difficulty on participants' perceived easiness of the task. By running a manipulation check, it can determine whether the intended manipulation of an independent variable (Level of difficulty) has been successful. If the manipulation check were successful, we can conclude that the manipulated variable has produced the desired effects on the participants or the experimental conditions (Hoewe, 2017). This thesis aimed to identify whether the proposed model, the SSAM model is a good fit for the collected data. To achieve this, structural equations modelling (SEM) was used. As suggested by Kline (1998) SEM can explain the maximum amount of variance from a specific model and help to understand the covariance or correlation patterns among a set of variables. Moreover, the analysis of SEM provides various statistical indices to evaluate the goodness of fit and overall quality of the SSAM model. These fit indices namely the Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), Root Mean Square Error of Approximation (RMSEA; Steiger & Lind, 1980), Akaike's Information Criteria (AIC; Akaike, 1974), and Bayesian Information Criteria (BIC; Stone, 1979), played a crucial role in assessing the adequacy of the model and its fitness to the observed data. CFI and TLI compare the hypothesis model to fit a baseline, while RMSEA measures the deviation from an ideal model. Moreover, the BIC and AIC measure model fit and future prediction against complexity. For a good fit, RMSEA should be below 0.06, and CFI and TLI should be larger than 0.95. Smaller AIC and BIC values indicate a better fit (Hu & Bentler, 1990; Mohammed et al., 2015).

By performing these analyses, we can determine whether the model is appropriate and whether it captures the underlying relationships in the data. As we intended to understand the patterns of different levels of difficulties in the handover tasks from SAE-L3 automated vehicles to the driver on situational trust, sleepiness, state anxiety, mental workload, and on performance. Therefore, the hypothesis was tested with the SEM approach with the lavvan packages. To examine the indirect effect of Situational Trust on Performance, and the SSAM model we applied the bootstrapping method. The bootstrap method (Razak et al., 2018) was used in this case in other to investigate whether the situational trust has a mediation impact. These effects were estimated using 5000 bootstrap samples and 95% of confidence interval.

Results

Descriptive Statistic on Variables and Level of Difficulty

Table 1

Mean and standard deviation of level of difficulty on variables.

Variables/ Level of difficulty	Easy	Moderate	Hard
	M (SD)	M (SD)	M (SD)
Situational Trust	4.69 (1.49)	4.83 (1.80)	4.72 (1.73)
Sleepiness	4.69 (1.18)	4.83 (1.55)	4.81 (1.77)
State anxiety	3.48 (0.90)	3.02 (1.24)	3.39 (1.23)
Mental workloads	31.46 (30.72)	17.83 (15.02)	52.72 (46.22)
Success	0.92 (0.27)	0.97 (0.96)	0.09 (0.15)

Regression Analysis on Trait Anxiety and State Anxiety

Then, a simple linear regression analysis was performed to determine whether trait anxiety affects state anxiety. The results of the regression analysis are presented in Table 1. The t-value of state anxiety is extremely significant ($t(34) = 4.829$). The Z-score analysis is conducted, but neither state anxiety nor trait anxiety shows any outliers.

Table 2

Regression Analysis of Trait Anxiety on State Anxiety

Variable	β	Std. Error	t-value	p-value
State Anxiety	3.88	0.70	4.83	<0.001

Trait Anxiety	-0.01	0.06	-0.14	0.893

Note. The confidence interval implied is $p < 0.05$.

Manipulation Check on the SSAM model

The result of the manipulation check is presented in Table 3. Based on the result, it seems that for the participants it was difficult to distinguish between the easy and moderate conditions ($p = 0.866$). As for hard conditions, it was significantly different from easy conditions ($p < 0.001$).

Table 3

Result of the manipulation check between the level of difficulty and participants' perceived easiness

Predictor	β	Std. Error	t-value	p-value
Easy	5.92	0.43	13.78	< 0.001
Moderate	-0.09	0.63	-0.15	0.886
Hard	-2.38	0.64	-3.75	<0.001

Note. The confidence interval implied is $p < 0.05$.

Hypothesis Testing

The evaluation of the fit between the observed data and the SSAM model (Figure 3) indicated strong fits (CFI = 1.000, TLI=1.000, RMSEA = 0.000). Moreover, the AIC and BIC values of 411.264 and 442.933, respectively, further suggest that the SSAM model provides a relatively good fit. However, it should be noted that the model may be overly fit for the data, as indicated by the perfect fit. The result of bootstrap analysis, as reported in Table 4, indicates that the independent variable (the level of difficulty), does not have a significant effect on the dependent variable, success (H3, β = -0.02, p = 0.744), nor does it significantly influence situational trust (H1, β = 0.05, p = 0.803).

Furthermore, situational trust does not emerge as a significant mediator in the SSAM model. Similarly, situational trust does not significantly impact the success of taking over (H2, β = 0.02, p = 0.859). The analysis of the effects of situational trust on the sub-factors (sleepiness, state anxiety, mental workload) revealed that: i) there is a significant effect of situational trust on mental workload (H10, β = -0.35, p = 0.011), and ii) on state anxiety (H12, β = -0.758, p <0.001). However, situational trust does not affect sleepiness (H11, β = -0.13, p = 0.454).

Additionally, the result revealed that a statistically significant correlation was observed between mental workload and state anxiety (H15, β = 0.32, p = 0.015). However, there is no significant correlation found between i) mental workload and sleepiness (H13, β =0.013, p = 0.896), and ii) sleepiness and state anxiety (H14, β = 0.11, p = 0.218). The visualisation of the hypothesis is presented in Figure 5.

Table 4

Results of Hypothesis

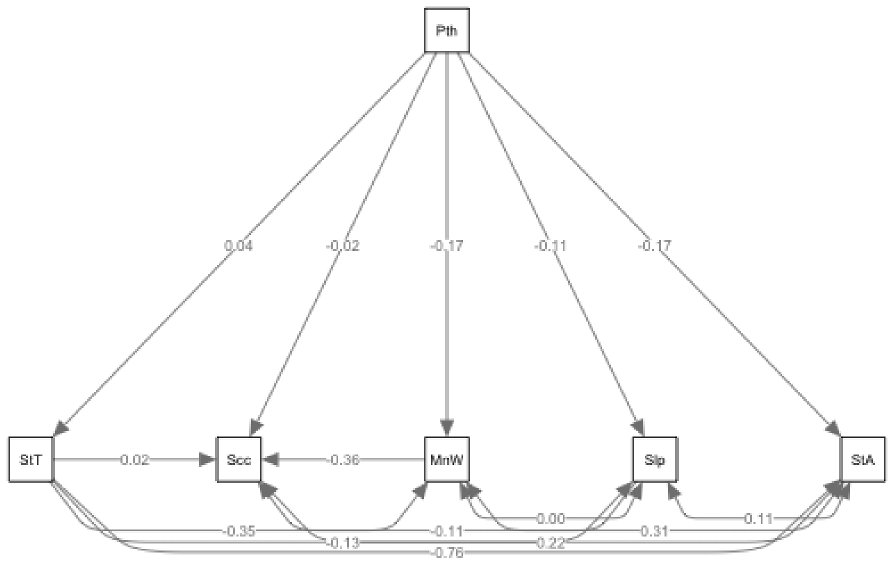
Hypothesis	Structural Relationship	β	S.E	p-value	Hypothesis supported?
H1	Level of difficulty impacted situational trust	0.05	0.21	0.803	
H2	Situational trust impacted the driver's takeover performance	0.02	0.12	0.085	
H3	The level of difficulty impacted the driver's takeover performance	-0.02	0.06	0.744	
H4	The level of difficulty impacted the driver's mental workload during the handover task	-0.09	0.13	0.497	

H5	The level of difficulty impacted the driver’s sleepiness during the handover task	-0.08	0.19	0.683	
H6	The level of difficulty impacted the driver's state anxiety during the handover task	-0.16	0.11	0.160	
H7	**Mental workload impacted the driver's successful takeover**	**-0.36**	**0.08**	**p<0.001**	**Yes**
H8	Sleepiness impacted the driver's successful takeover	-0.11	0.08	0.147	
H9	**State anxiety impacted the driver's successful takeover**	**0.22**	**0.11**	**0.054**	**Yes**
H10	**Situational trust impacted the driver's mental workload during the handover task**	**-0.35**	**0.14**	**0.011**	**Yes**
H11	Situational trust impacted the driver's sleepiness during the handover task	-0.13	0.18	0.454	
H12	**Situational trust impacted the driver's state anxiety during the handover task**	**-0.76**	**0.11**	**p<0.001**	**Yes**
H13	Mental workload impacted sleepiness	0.01	0.10	0.896	
H14	Sleepiness impacted state anxiety	0.11	0.09	0.218	
H15	**Mental workload impacted state anxiety**	**0.32**	**0.13**	**0.015**	**Yes**

Note. The confidence interval implied is $p < 0.05$.

Figure 5

The SEM visual output of the SSAM model



Note. "Pth" refers to the level of difficulty in the handover task. “StT" refers to situational trust. "Scc" refers to the success in taking over performance. "MnW” refers to mental workload. “Slp” refers to sleepiness. “StA” refers to state anxiety.

Exploring an alternative model

The manipulation checks on the SSAM model have suggested no difference in terms of the success of taking over during the easy and moderate/medium difficulty conditions. Therefore, we collapse the data of the easy and moderate conditions. This resulted in an alternative model (Model 2) to test our hypotheses.

Manipulation check of Model 2

In Table 5, the manipulation checks on Model 2 show that under the merged condition, more variations were found between Easy+Moderate ($p < 0.001$) and Hard ($p < 0.001$).

Table 5

Result of the manipulation check between the level of difficulty and participants' perceived easiness for model 2

Predictor	β	Std.Error	t-value	p-value
Easy + Moderate	5.88	0.31	19.24	<0.001
Hard	-2.34	0.55	-4.22	< 0.001

Note. The confidence interval implied is p <0.05.

Similarly, to the previous model, Model 2 seems to strongly fit (even to overfit) the data (CFI = 1.000, TLI=1.000, RMSEA = 0.000). However, the AIC and BIC values of 344.43 and 376.09, respectively. These results suggest that the alternative model fits the data respectively well.

In Table 6, the bootstrapped result of the alternative model is presented. The result revealed that the independent variable (level of difficulty), has a significant effect on only one dependent variable, success (H3, $\beta = -0.33$, $p =<0.001$). However, this effect is not found on situational trust (H1, $\beta = -0.06$, $p = 0.800$). Moreover, the Level of difficulty showed a significant impact on mental workload (H4, $\beta = 0.69$, $p =<0.001$). No significant impacts were found in the relationship between the level of difficulty and the driver's sleepiness during the task (H5, $\beta = 0.079$, $p = 0.70$).

Furthermore, the sub-factors significantly impacted success in takeover performance (H7, $\beta = -0.26$, $p =<0.001$; H8, $\beta = -0.09$, $p = 0.013$; H9, $\beta = 0.08$, $p = 0.035$). The analysis of the effects of situational trust on sleepiness, state anxiety, and mental workload revealed that: i) there is a significant effect of situational trust on sleepiness (H11, $\beta = -0.30$ $p = 0.004$), and ii) on state anxiety (H12, β = -0.40, p <0.001). However, situational trust does not affect mental workload (H10, β = -0.08, p = 0.341).

Additionally, the result suggested only mental workload has an impact on state anxiety (H15, β = 0.41, p = 0.001). However, there is no effect found between mental workload and sleepiness (H13, β = -0.13, p = 0.136) and on sleepiness and state anxiety (H14, β = -0.50, p = 0.581). View Figure 6 for an overview of the alternative model.

Table 5

Results of Hypothesis on an alternative model (Model 2)

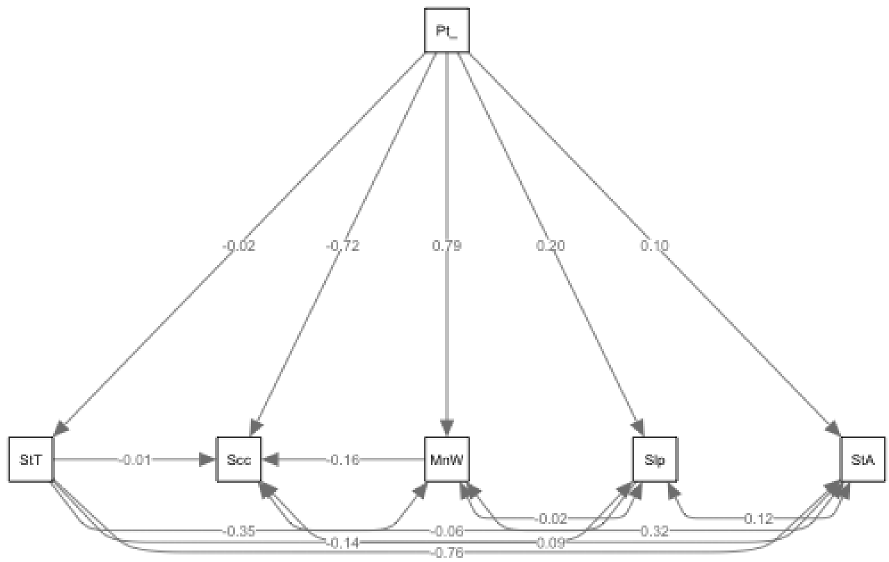
Hypothesis	Structural Relationship	β	S.E	p-value	Hypothesis supported?
H1	Level of difficulty impacted situational trust	-0.060	0.23	0.800	
H2	Situational trust impacted the driver's takeover performance	-0.01	0.03	0.675	
H3	**The level of difficulty impacted the driver's takeover performance**	**-0.33**	**0.07**	**p<0.001**	**Yes**
H4	**The level of difficulty impacted the driver's mental workload during the handover task**	**0.70**	**0.18**	**p<0.001**	**Yes**
H5	The level of difficulty impacted the driver's sleepiness during the handover task	0.080	0.20	0.695	
H6	The level of difficulty impacted the driver's state anxiety during the handover task	0.22	0.20	0.276	
H7	**Mental workload impacted the driver's successful takeover**	**-0.26**	**0.03**	**p<0.001**	**Yes**
H8	**Sleepiness impacted the driver's successful takeover**	**-0.09**	**0.04**	**0.013**	**Yes**
H9	**State anxiety impacted the**	**0.08**	**0.04**	**0.035**	**Yes**

	driver's successful takeover				
H10	Situational trust impacted the driver's mental workload during the handover task	-0.08	0.08	0.341	
H11	**Situational trust impacted the driver's sleepiness during the handover task**	**-0.30**	**0.10**	**0.004**	**Yes**
H12	**Situational trust impacted the driver's state anxiety during the handover task**	**-0.40**	**0.10**	**p<0.001**	**Yes**
H13	Mental workload impacted sleepiness	-0.13	0.09	0.136	
H14	Sleepiness impacted state anxiety	-0.05	0.09	0.581	
H15	**Mental workload impacted state anxiety**	**0.41**	**0.09**	**p<0.001**	**Yes**

Note. The confidence interval implied is $p < 0.05$.

Figure 6

The SEM visual output of the alternative model (Model)2



Note. "Pth" refers to the level of difficulty in the handover task. "StT" refers to situational trust. "Scc" refers to the success in taking over performance. "MnW" refers to mental workload. "Slp" refers to sleepiness. "StA" refers to state anxiety.

Discussion

The main objective of this paper was to examine the impact of people's situational trust on successfully taking over the control of Level 4 autonomous cars at different levels of difficulties of the handover task. Additionally, we also investigated whether situational trust influences the sub-factors (Sleepiness, State Anxiety, and Mental workload) and whether these factors affect the success of taking over. To investigate these complex relationships, the SSAM model (Figure 3) was proposed as a theoretical framework. To achieve that, we employed a combination of virtual reality technology and a driving simulator to create a realistic driving environment. Furthermore, a questionnaire was developed, consisting of multiple items that assessed participants' situational trust and the sub-factors during the handover task.

The regression results in Table 1, suggested that there is no significant effect between trait anxiety and state anxiety. In this study, it was determined that the driver's trait anxiety did not affect the driver's state anxiety during the performer's takeover. However, the t-value for state anxiety seemed extreme in this study. We suspect that such extreme values result from potential outliers in the driver's state anxiety, therefore we run a z-score analysis to identify them. The Z-score analysis does not report any outliers within the data.

The result of this study provided support for H7, indicating a significant impact of mental workload on the success of takeover performances in autonomous driving. These findings suggest that the level of mental workload experienced by drivers during the handover task directly influences their ability to successfully assume control of the automated vehicles. The supported hypothesis suggested a negative relationship between mental workload and driver's takeover performance, implying that drivers are more likely to achieve successful takeovers when the handover task requires a lower mental workload. Our result aligned with the study of

Bueno et al (2016). They observed the same negative effect between mental workload, non-driving related tasks and takeover performances of drivers in SAE-L3 autonomy. This alignment highlights the importance of considering mental workload as a critical factor to better understand takeover performances in high-level autonomous driving.

Next, the findings of this study confirmed H9, providing evidence for a significant impact of state anxiety on the success of takeover performance in the context of Level 3 automated vehicles. During the handover task, higher levels of state anxiety were associated with poorer performance in terms of drivers' ability to assume control of automated vehicles. Therefore, the accepted hypothesis suggests a negative relationship between state anxiety and takeover success. This explained that a higher level of state anxiety may affect the driver’s ability to regain control of the vehicle from an automated system. Such a negative relationship between state anxiety and takeover performances has not been identified in previous research. However, our research provides evidence to support this relationship. Therefore, more research should be done to further investigate the role of state anxiety on takeover performance in high levels of autonomous driving.

H10, predicting that situational trust affects mental workload was also found to be significant. Based on the evidence provided above, situational trust negatively impacted the driver’s mental workload during the handover task. This could further explain that higher levels of situational trust in the SAE-L3 automated system are associated with the lower mental effort required by drivers during the handover task. These findings align with previous studies conducted by Clement et al. (2022) and Yousfi et al. (2022). In a similar study, Clement et al. (2022) observed that trust is related to mental workload, with higher levels of trust correlated with lower levels of workload. Furthermore, they highlighted the importance of considering the driver's age and experience when examining the effects of trust on mental workload. We addressed this aspect in our study by recruiting participants within the same age group (20 to 25 years old), thereby enhancing validity. Our results are also consistent with those of Yousfi et al. (2022), who reported that higher levels of trust were associated with lower mental effort. Likewise, Du et al. (2019) reported that increased trust results in a reduction of mental workload. Thus, our study contributes to the existing literature by demonstrating a significant effect of situational trust on mental workload. This reinforces the conclusions drawn by Clement et al. (2022) and Yousfi et al. (2022). Based on this result we can conclude that drivers who have more confidence and rely on the automated system are likely to experience less cognitive load when assuming control, which may improve their overall performance.

Moreover, the twelfth hypothesis indicated a significant impact of situational trust on drivers' state anxiety during the handover task in the context of SAE-L3 automated driving. As proposed in the H12, situational trust negatively affected the driver's state anxiety during the context of the handover task. This finding suggested that drivers who have a greater level of trust in the SAE-L3 automated system are more likely to exhibit lower levels of anxiety during the handover task. Similar findings were reported by Lu et al. (2020) regarding lower levels of anxiety among drivers that trusted automated systems. Based on the converging results across studies, the observed relationship is more valid and highlights the importance of trust in reducing driver anxiety during critical moments of transitioning control from the automated system to the driver. Our study, by utilizing the insights provided by Lu et al. (2020), contributes to a growing body of research that acknowledges trust as a crucial factor in influencing the emotional state of the driver during handover scenarios. By focusing on the role of trust in reducing anxiety, future research and system design efforts can focus on ways to build trust. Therefore, establishing trust in the automated system can reduce psychological stress and anxiety related to the transfer of control between the driver and the system. Developing trust in automated systems is a crucial step to enhancing driver comfort and confidence during handover scenarios, ultimately enhancing performance and safety in autonomous driving scenarios.

Lastly, H15 reflected the relationship between mental workload and state anxiety was found significant. The expectation of the idea that mental workload has positive effects on state anxiety is confirmed by the provided evidence in the result section. This finding suggested that drivers who experience heightened mental effort during the handover task are more likely to experience higher levels of anxiety. This highlights the interplay between mental workload and state anxiety, indicating that cognitive demands during handover tasks have a direct impact on the driver's psychological state. This result indicated the importance of managing mental workload during autonomous driving scenarios to reduce the negative effects of state anxiety.

Contrary to expectation, the first hypothesis, which posited a negative impact of the level of difficulty of the handover task on the driver's situational trust in Level 3 autonomous driving scenarios, yielded insignificant results. This finding suggested that the level of difficulty in the handover task does not exert a significant influence on the driver's situational trust in the Level 3 automation system. Despite our initial expectation that greater task difficulty would diminish drivers' trust in automated vehicles, the empirical evidence does not support this hypothesis. In contrast to Yousfi et al. (2022), our study did not reveal a significant relationship between the level of difficulty and situational trust. We found no significant association between these variables, contrary to their findings. While our study did not find significant correlations between the level of difficulty of the handover task and the variables of interest, this does not invalidate the research of Yousfi et al. (2022). Moreover, our study's non-significant findings regarding the impact of handover task difficulty on the variables of interest do not diminish the importance of considering it. It is worth mentioning that as suggested by the manipulation check, the designed window of opportunity in easy condition and moderate conditions has a limited variation, this may have constrained the observed relationship. Thus, we speculate that these constraints in the experimental design may have contributed to the non-significant result in this relationship.

Despite our expectations, the findings of this study revealed that situational trust does not significantly impact a driver's success in takeover performance (H2). A positive relationship between situational trust and takeover performance was not supported by the data analysis of hypothesis 2. Our initial hypothesis suggested that higher levels of situational trust in the SAE-L3 autonomous driving context would be associated with a higher likelihood of successful takeovers. However, the results indicate that situational trust alone may not be a determining factor in a driver's takeover performance.

The findings of this study do not support hypothesis 3, which proposed a negative impact of the level of difficulty on the driver’s takeover performance. We expected that, as the level of difficulty of the handover task increased, drivers would experience a decrease in their ability to successfully assume control of the Level 3 automated vehicle. However, no significant relationship between these variables is not revealed. This suggested that the level of difficulty of the handover task may not be the sole determinant of the driver’s takeover performance. The limited variation in the timeframe structure for each stage of the transfer of tasks, as discussed previously, could explain the lack of significant effect of difficulty on the success of TOR.

We speculate that similar reasons could account for the lack of significant results regarding the effects of the level of difficulty on mental workload, sleepiness or on state anxiety. These results imply that the level of difficulty in the handover task may not directly influence these factors as hypothesised. Moreover, these results further supported the fact that the design timeframe for takeover performance holds limited variations.

In the eighth hypothesis, sleepiness was hypothesized to negatively affect drivers' takeover performance during the takeover task. According to H8, the increased level of driver fatigue during the handover task predicts a lower takeover success rate. It turned out, however, that the analysis did not support this expectation. The findings indicate that fatigue levels of drivers during the handover task may not directly influence takeover performance as hypothesised. This suggested a need for further investigation into the complex interplay between sleepiness and takeover performances during handover tasks.

Accordingly, we anticipated that situational trust would affect the driver's sleepiness during the handover task (H11) in accordance with Kundinger et al. (2019). In the context of automated vehicles, Kundinger et al. (2019) argue that drivers who trust automation are more likely to be sleepy. This expectation was not met by our results. There was no evidence to support the hypothesis that sleepiness impacts takeover performance positively. There may be a more complex relationship between situational trust and driver sleepiness during handover than first thought. However, it is crucial to note that the incongruity between our findings and the study by Kundinger et al. (2019) does not invalidate their research. This insignificant result can be explained by the differences in study design and methodology that may account for the disparate results. For instance, Kundinger et al. (2019) focused on general trust in automation, whereas we focused on situational trust. Moreover, we emphasize Level 3 automation whereas Kundinger et al. (2019), placed focus on Level 2 automation, the difference between the automation level can provide a further explanation of the contradicted research result. The lack of evidence to support a positive effect of sleepiness on driver's takeover performance in our study does not necessarily negate Kundinger et al.'s (2019) findings. Rather, it highlights the need for further research and a nuanced understanding of the complex relationship between situational trust, sleepiness, and takeover performance.

Lastly, the present data cannot support the assumption that drivers would feel less sleepy when facing a higher mental workload during the handover task (H13), and that state anxiety would affect driver sleepiness during the handover task (H14).

Overall, this study indicated that situational trust did not act as a mediator in the relationship between the SSAM model variables. This conclusion was supported by the rejection of hypotheses two and eleven. While hypotheses ten and twelve were accepted, suggesting that situational trust had an impact on driver sleepiness and state anxiety during the handover task. However, it is critical to note that these two supported hypotheses alone do not provide sufficient evidence to establish situational trust as a mediator.

There may be other influential factors that were not considered in the research design due to the non-mediation of situational trust in this study. Aside from the sub-factors (sleepiness, state anxiety), other variables may also affect takeover performance. Therefore, further research is needed to identify and investigate these possible mediating factors. However, the findings did highlight the importance of situational trust as a significant variable in the context of takeover performance and sub-factors. In light of this, situational trust is a relevant factor to consider, although its precise role and influence may not be as simple as first thought.

Another noteworthy point in this discussion section is that the proposed SSAM model overfits data. Due to overfitting, the parameters of the model were unduly influenced by the characteristics and idiosyncrasies of the data sample used for bootstrapping, which could have compromised the accuracy of the model to reflect the underlying patterns and relationships in the dataset (Montesinos-López et al., 2022). As a result, the fit of the model becomes excessively tailored to the complexities of the observed data, raising concerns about generalizability beyond the sample. It is important to consider the validity and stability of the estimated parameters in the context of the SSAM model upon discovering overfitting, as well as the model's ability to accurately depict the underlying processes and dynamics associated with the phenomenon under study. Certain relationships proposed in this study could be overestimated, potentially resulting in biased conclusions due to an overestimation of their magnitude and significance. The overfitted SSAM model does not render the entire model useless but rather emphasizes the need for refinement and validation to improve its performance and applicability. It serves as a starting point for identifying associations and generating hypotheses, guiding the development of more robust models in future studies.

The exploratory analysis of Model 2 suggested that this model is less efficient and effective (AIC = 344.427 and BIC = 376.097) than the previous model (AIC = 411.264 and BIC = 442.933). Nevertheless, Model 2 suggests a significant effect for eight of the proposed hypotheses compared to the five hypotheses proposed from the original models.

Therefore, model 2 in addition to the original model suggests that the level of difficulty affects the driver's takeover performance (H3). According to the third hypothesis, the higher level of task difficulty predicts a lower chance of successfully taking over from the automated vehicles. This explained that the previous insignificant result is due to the limited variation in the designed handover task. This highlights the interplay between the level of difficulty and the success of the driver's takeover performance during the handover task. So far, no study has been found to examine such an effect we proposed, therefore based on the result we achieve, we conclude a negative relationship between the level of difficulty and the driver's success in taking over.

Similarly, the fourth hypothesis was also found to be significant after the merging of two conditions. As proposed in H4, the mental workload will be negatively affected by the level of difficulty during the handover task. This result indicated that more mental workload is required when the perceived level of difficulty was increased during the handover task. Hypothesis four was proposed as an explorative expectation between mental workload and level of difficulty. No prior studies have investigated this relationship. Therefore, based on the result of this study, we concluded that there is a negative relationship found between the level of difficulty and mental workload.

The result of the eighth hypothesis has become significant after the merging of two conditions. According to the results, sleepiness has a negative effect on the driver's takeover performance. Thus, fatigued drivers are less likely to successfully takes control of the handover task. Moreover, the findings supported the eleventh hypothesis, indicating that situational trust positively affects sleepiness. This aligns with Kundinger et al's study. (2019). Based on the author's analysis, the sleepiness of the driver affects how well the automated vehicle takes over from the driver. Differing from Kundinger et al. (2019), our focus is on SAE-L3 automation, therefore this result does not just confirm our expectation of these variables, but most importantly, it added support to the role of sleepiness in both levels of automation.

Interestingly, Model 2 contradicted the SSAM model's significant results. Specifically, hypothesis ten, which predicted a positive impact of situational trust on mental workload, was not supported in Model 2. This change in significance may be attributed to the manipulation of the level of difficulty in the handover task.

By merging the conditions in Model 2, the variation in the level of difficulty was increased. This merging likely weakened the relationship between situational trust and mental workload, leading to an insignificant result. Based on these findings, situational trust affects mental workload significantly depending on the level of difficulty. In addition, this study indicates that situational trust might not be a sufficient mediator. Further research is needed to determine whether other factors or variables are influencing the relationship between situational trust and mental workload.

Limitations

The present study represents an initial endeavour to construct models encompassing situational trust, sleepiness, state anxiety, mental workload and handover performance. Consequently, this study encountered various limitations that might have influenced the study outcomes. Firstly, the study encountered a limitation in conducting the present research objectively. Technical difficulties hindered the design of the driving simulator program, preventing the detection and recording of participants' handover performance. As a result, the observation method was employed, requiring the presence of at least two researchers throughout the experiment. This was to assess whether participants successfully or unsuccessfully completed the handover task. These limitations introduce challenges in determining whether participants' last-millisecond reactions should be classified as successes or failures. Likewise, our inability to provide instant feedback through the VR headset due to technical difficulties also limited our ability to provide instant feedback. We had to rely on verbal feedback to inform participants if the handover task was completed successfully. Considering the verbal feedback participants received, this deviation from the original plan caused a certain level of stress for them. Furthermore, the use of subjective questionnaires to collect participants' assessments of situational trust and sub-factors may have introduced subjective bias to the results. There is a possibility that participants' perceptions and interpretations of the questions could differ, potentially affecting the accuracy and reliability of the collected data.

Secondly, the driving simulation program design had limitations that affected the participants' experience. As a result of the automated vehicle's inability to drive smoothly and the VR headset's shaky vision, some participants experienced nausea. Over time, however, participants reported becoming more accustomed to the shakiness as they progressed through the scenarios. However, it is important to note that several participants reported that the level of shakiness experienced during the driving simulation did have an impact on their perceived situational trust towards the automated vehicles. Extra waypoints were added to prevent collisions or veering off the road, consequently, the automated vehicles failed to stop automatically at the end of each scenario. Furthermore, we experienced frequent malfunctions with Unity, (the program used to design the driving simulators and VR environment) during the experiment. These malfunctions resulted in system crashes, causing interruptions and pauses in the running of the programs. As a result, the flow of the participants' driving experience in the artificial environment was significantly disrupted. Due to this, one participant had to redo a scenario during the experiment, which could have potentially influenced the result of the study.

A third limitation of the study is that only six participants were recruited, limiting the generalizability of the results. As a result of the small sample size, it may be difficult to draw broader conclusions and generalize the findings to a larger population. Additionally, the requirement that participants possess a valid driving license and have driving experience further restricted the participant pool, potentially limiting diversity and representativeness.

Fourthly, the driving environment was not designed to replicate the realistic function of side mirrors and rear-view mirrors in this study. Real-life driving situations are enhanced by side mirrors, which provide a wider view of traffic and the surrounding environment, whereas rear-view mirrors give drivers a better view of behind-the-vehicle traffic and objects. It was found, however, that the view in both side mirrors and rear-view mirrors were identical and did not change with car movement. Participants' perceptions and behaviour during the handover task may have been affected by this discrepancy from real-world conditions. Because of this discrepancy, participants were confused by the driving environment design. Mirrors with inaccuracies have caused participants to feel disoriented and uncertain about surrounding cars and objects. During the handover task, this confusion could have compromised ecological validity by influencing their level of situational trust and handover performance.

Another limitation of this study is the potentially illogical nature of the handover task presented to participants. Participants were given predetermined timeframes to take over vehicle control. This was without considering the absence of a clear reason or obstacle necessitating the handover, such as an obstruction on the road. Consequently, participants found the task confusing and unnatural, leading to discomfort and a potential impact on their perceived situational trust in automated vehicles. The absence of a realistic context for the handover task may have hindered participants' engagement and compromised the ecological validity of the study.

Finally, the limited variation in the level of difficulty during a handover exercise is one limitation of this study. The manipulation checks indicated that there was a small difference between easy and moderate conditions. In this absence of differentiation, participants may not have been able to fully evaluate and take into account the differences in conditions. Therefore, by including a broader range of difficulties in the task to be handed over, further studies should seek to tackle this limitation. This will allow a more comprehensive examination to be made of the impact on variables of interest due to handover task difficulty and provide greater insight into the phenomenon under investigation.

Recommendations

Based on the limitations encountered in this study, there are several recommendations for future research. Firstly, it is crucial to enhance the level of objectivity in measuring handover performance. This can be achieved by designing the driving simulator program to accurately detect and record the participants' handover performances during the HOR task. By utilising objective measurement methods, such as data logging or automated performance assessment algorithms, the reliance on subjective observations can be minimised, reducing potential biases. Additionally, we propose the inclusion of physiological measurements, such as Electrodermal Response (EDR) and heart rate variability (HRV), as objective indicators of participants' physiological states. Although originally intended in our study, technical difficulties hindered the collection of these data. Future researchers are encouraged to incorporate wearable devices, such as the Shimmer 3 GSR+, to capture EDR and HRV data, providing valuable insights into participants' physiological responses during handover scenarios. These physiological measurements can provide a more comprehensive understanding of the participants' psychophysiological states and their relationship with situational trust and handover performance. Moreover, it would be ideal to provide the participants with the outcome of their handover performance directly through the VR headset. This would eliminate the need for verbal feedback, which has the potential to induce stress and affect participant performance.

Furthermore, it is essential to carefully design the driving environment in a way that ensures the smooth operation of the simulated vehicle. This includes avoiding the addition of unnecessary waypoints that may disrupt the automated driving experience. However, it is crucial to strike a balance as removing these extra waypoints can potentially lead to simulated vehicles crashing or running off the road at the end of the scenario. Future research should focus on finding an optimal configuration that maintains a realistic driving experience while minimising technical issues and inconsistencies in the simulation.

Additionally, future studies should aim to broaden the criteria for participant recruitment to enhance the generalizability of the results. In our study, we specifically focused on participants with valid driving licenses and driving experience, which inadvertently excluded individuals without driving licenses. This limitation is unfortunate, considering that our study explored the context of SAE-level 3 automation. Therefore, future research should strive to include participants from diverse backgrounds and demographics to capture a broader range of perspectives and experiences. Moreover, our study primarily relied on recruiting university students as participants, which may introduce sampling bias and limit the generalizability of the findings to a specific population. To overcome this limitation, future studies should aim to recruit participants from different age groups and occupational backgrounds. Including participants from various professions, such as healthcare workers, office employees, or retirees, can provide valuable insights into the impact of situational trust and handover performance across different occupational contexts.

Fourthly, future studies should aim to design the driving simulator with realistic functionality of side mirrors and rear-view mirrors. This ensures that the view of mirrors corresponds to the movement of the car, providing participants with an accurate representation of their driving surroundings. With this incorporation of realistic mirror functionality, future studies can enhance the ecological validity of the driving environment and minimise participants' confusion.

In addition, we recommend future studies consider providing a more realistic and meaningful context for the handover task. This can be achieved by incorporating situations that warrant handover, such as a tire blocking the road or taking the left exist to avoid traffic jams. By contextualising the handover tasks within real-world scenarios, participants can better understand the purpose and relevance of the TOR task and reaction in a more efficient manner.

Last but not least, we recommend increasing the variability in the level of difficulty during handover tasks in future studies. By doing so, participants would be able to more accurately reflect on and observe differences between conditions. With a wider range of difficulty levels, researchers can gain a deeper understanding of how difficulty affects various variables.

Conclusion

In the present study, we attempted to: i) gain a better understanding of the relationship between situational trust, level of difficulty, and handover performance through this study, and ii) explore the role of sub-factors (sleepiness, state anxiety and mental workload) in the process of handover to humans. The theoretical model we tested seems promising, despite looking at our results situational trust does not serve as a mediating factor in the relationships as intended, it still holds considerable value as an influential element. While the direct impact of situational trust on the ability to take over in difficult situations was not clarified in the present pilot study, gathering data from a larger group of participants after a redesign of the experiment as recommended above, could lead to better insights about the role of trust, and its connection with sub-factors in the context of taking over the command while driving an autonomous car. In fact, our results highlighted the potential interaction between situational trust and the sub-factors. This research offers the preliminary formalization and testing of a theoretical model that holds together trust, sleepiness, workload, and anxiety with handover complexity and success in taking over autonomous vehicles. This model could be the starting point for future investigation on the role of situational trust, sleepiness, state anxiety and, mental workload in autonomous driving.

References

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723. https://doi.org/10.1109/tac.1974.1100705

Åkerstedt, T., & Gillberg, M. (1990). Subjective and Objective Sleepiness in the Active Individual. *International Journal of Neuroscience*, *52*(1–2), 29–37. https://doi.org/10.3109/00207459008994241

Avetisian, L., Ayoub, J., & Zhou, F. (2022). Anticipated emotions associated with trust in autonomous vehicles. *Proceedings of the Human Factors and Ergonomics Society . . . Annual Meeting*, *66*(1), 199–203. https://doi.org/10.1177/1071181322661002

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Cai, A. W., Manousakis, J. E., Singh, B., Francis-Pester, E., Kuo, J., Jeppe, K. J., Rajaratnam, S. M. W., Lenné, M. G., Howard, M. E., & Anderson, C. (2023). Subjective awareness of sleepiness while driving in younger and older adults. *Journal of Sleep Research*. https://doi.org/10.1111/jsr.13933

Chen, W., Sawaragi, T., & Horiguchi, Y. (2019). Measurement of Driver's Mental Workload in Partial Autonomous Driving. *ResearchGate*. https://www.researchgate.net/publication/336363965_Measurement_of_Driver's_Mental_Workload_in_Partial_Autonomous_Driving

Clement, P., Veledar, O., Konczol, C., Danzinger, H., Posch, M., Eichberger, A., & Macher, G. (2022). Enhancing Acceptance and Trust in Automated Driving through Virtual Experience on a Driving Simulator. *Energies*, *15*(3), 781. https://doi.org/10.3390/en15030781

Du, N., Haspiel, J., Zhang, Q., Tilbury, D. M., Pradhan, A. K., Yang, X., & Robert, L. P. (2019). Look Who's Talking Now: Implications of AV's Explanations on Driver's Trust, AV Preference, Anxiety and Mental Workload. *Transportation Research Part C-emerging Technologies*. https://doi.org/10.1016/j.trc.2019.05.025

Ferraro, J., & Mouloua, M. (2022). Assessing How Driving Self-Efficacy Influences Situational Trust in Driver Assist Technologies. *Proceedings of the Human Factors and Ergonomics Society . . . Annual Meeting*, *66*(1), 1260–1264. https://doi.org/10.1177/1071181322661476

Hoewe, J. (2017). Manipulation Check. *The International Encyclopedia of Communication Research Methods*, 1–5. https://doi.org/10.1002/9781118901731.iecrm0135

Hoff, K. A., & Bashir, M. (2015). Trust in Automation. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jin, M., Lu, G., Chen, F., Shi, X., Tan, H., & Zhai, J. (2021). Modeling takeover behavior in level 3 automated driving via a structural equation model: Considering the mediating role of trust. *Accident Analysis & Prevention*, *157*, 106156. https://doi.org/10.1016/j.aap.2021.106156

Kim, H., Kim, W. S., Kim, J., Lee, S. H., Yoon, D., Kwon, O. J., & Park, C. S. (2022). Study on the Take-over Performance of Level 3 Autonomous Vehicles Based on Subjective Driving Tendency Questionnaires and Machine Learning Methods. *Etri Journal*, *45*(1), 75–92. https://doi.org/10.4218/etrij.2021-0241

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

Kourtesis, P., Linnell, J., Amir, R., Argelaguet, F., & MacPherson, S. E. (2023). Cybersickness in Virtual Reality Questionnaire (CSQ-VR): A Validation and Comparison against SSQ and VRSQ. *Virtual Worlds 2023, 2(1)*, *2*(1), 16–35. https://doi.org/10.3390/virtualworlds2010002

Kundinger, T., Wintersberger, P., & Riener, A. (2019). *(Over)Trust in Automated Driving*. https://doi.org/10.1145/3290607.3312869

Lu, Y., Yi, B., Song, X., Zhao, S., Wang, J., & Cao, H. (2022). Can we adapt to highly automated vehicles as passengers? The mediating effect of trust and situational awareness on role adaption moderated by automated driving style. *Transportation Research Part F-traffic Psychology and Behaviour*, *90*, 269–286. https://doi.org/10.1016/j.trf.2022.08.011

Mayer, R., Davis, J., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, *20*(3), 709. https://doi.org/10.2307/258792

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

Mohammed, E. A., Naugler, C., & Far, B. H. (2015). Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. In *Elsevier eBooks* (pp. 577–602). https://doi.org/10.1016/b978-0-12-802508-6.00032-6

Muckler, F. A., & Seven, S. A. (1992). Selecting Performance Measures: "Objective" versus "Subjective" Measurement. *Human Factors*, *34*(4), 441–455. https://doi.org/10.1177/001872089203400406

Muir, B. M. (1994). Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics*, *37*(11), 1905–1922. https://doi.org/10.1080/00140139408964957

Razak, N. a. A., Zamzuri, Z. H., & Suradi, N. R. M. (2018). Bootstrapping technique in structural equation modeling: a Monte Carlo study. *Journal of Physics*, *1132*, 012072. https://doi.org/10.1088/1742-6596/1132/1/012072

SAE. (2021, May 3). SAE International. Retrieved March 7, 2023, from https://www.sae.org/blog/sae-j3016-update

Soares, S., Lobo, A., Ferreira, S., Cunha, L., & Couto, A. (2021). Takeover performance evaluation using driving simulation: a systematic review and meta-analysis. *European Transport Research Review*, *13*(1). https://doi.org/10.1186/s12544-021-00505-2

Spielberger, C. D., & Smith, L. (1966). Anxiety (drive), stress, and serial-position effects in serial-verbal learning. *Journal of Experimental Psychology*, *72*(4), 589–595. https://doi.org/10.1037/h0023769

Steiger, J. H., & Lind, J. C. (1980). Statistically Based Tests for the Number of Common Factors. Paper Presented at the Psychometric Society Annual Meeting, Iowa City, IA.

Stone, M. B. (1979). Comments on Model Selection Criteria of Akaike and Schwarz. *Journal of the Royal Statistical Society Series B-methodological*, *41*(2), 276–278. https://doi.org/10.1111/j.2517-6161.1979.tb01084.x

Tucker, L. R., & Lewis, C. E. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/bf02291170

Yousfi, E., Malin, S., Halit, L., Roger, S., & Dogan, E. (2021). *Driver experience and safety during manual intervention in a simulated automated vehicle*. https://doi.org/10.1145/3452853.3452888

Zheng, H., Mason, J. R., Classen, S., & Giang, W. C. W. (2023). Pilot study: Effect of roles and responsibility training on driver's use of adaptive cruise control between younger and older adults. *Transportation Research Part F: Traffic Psychology and Behaviour*, *94*, 53–66. https://doi.org/10.1016/j.trf.2023.01.023

Zijlstra, Fred. (1993). Efficiency in Work Behavior: A Design Approach for Modern Tools

Appendix A
Experiment equipment and set up.

This appendix consists of the experiment material that was used for this driving simulator study.

Figure A1
Driving simulator

Note. model of the driving simulator is Logitech G920 Driving Force combine with the next level racing chair.

Figure A2
Steering wheel of the driving simulator

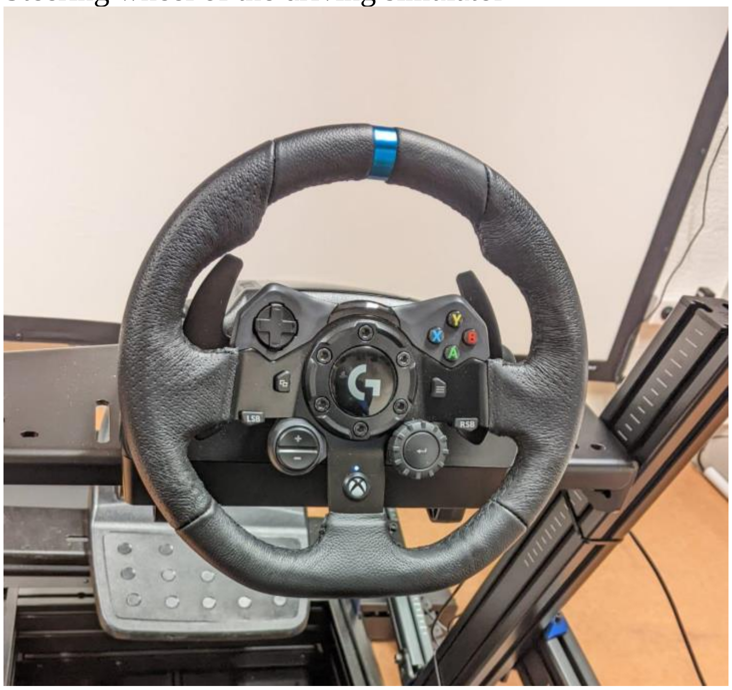


Figure A3
Pedals of driving simulator

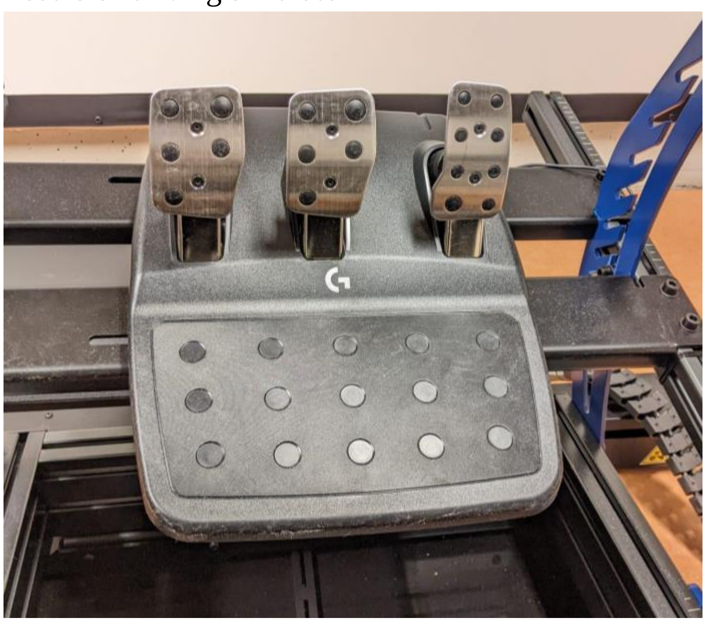


Figure A4
Virtual Reality Headset

Note. the model of the VR headset is Varjo XR-3

Figure A5
Full Experiment setup with participant

Figure A6

Unity program

Figure A7

Model of the virtual reality automated vehicle

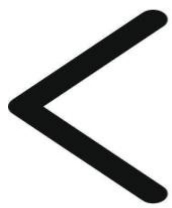
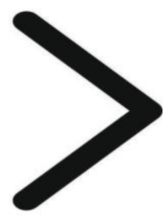
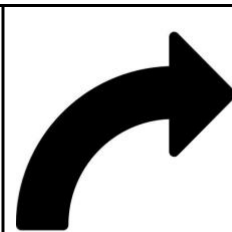
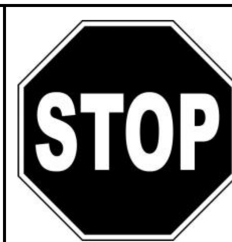
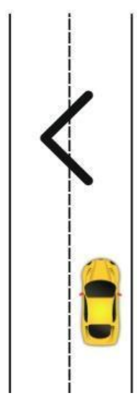
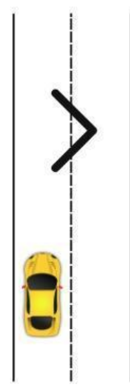
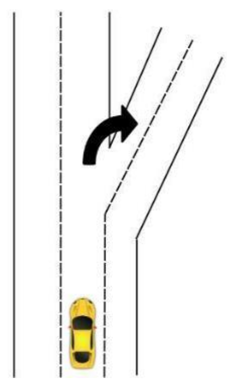
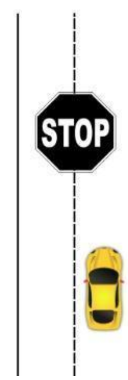
Note. Volvo XC40 recharge

Appendix B

Materials

Table B1

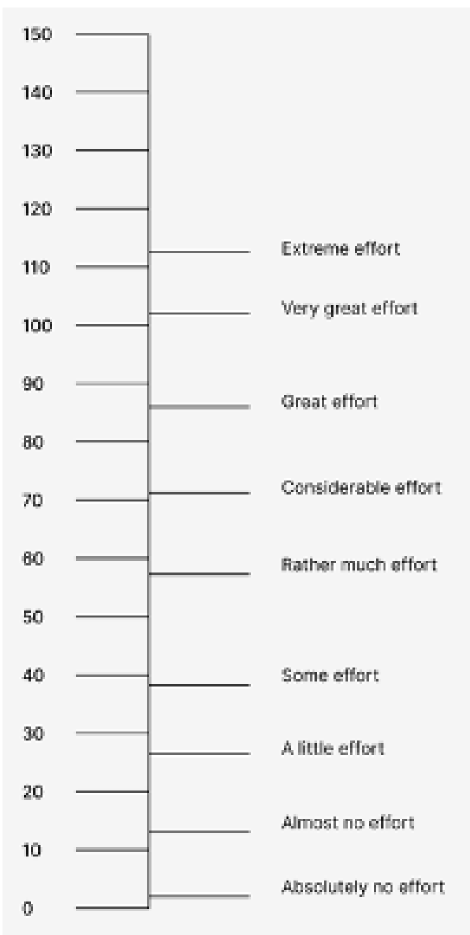
Scenarios condition and TOR instructions

TOR sign				
Instruction	Change to left lane	Change to right lane	Take the exit to the right	Brake
Diagram				
Window of opportunity (Levels of difficulty / seconds)	Training/ 25 sec Easy/ 5 sec Moderate /3 sec Hard / 1 sec	Training/ 25 sec Easy/ 5 sec Moderate /3 sec Hard / 1 sec	Training/ 25 sec Easy/ 5 sec Moderate /3 sec Hard / 1 sec	Training/ 25 sec Easy/ 5 sec Moderate /3 sec Hard / 1 sec
Desired action	Participants steer the steering wheel to the left	Participants steer the steering wheel to the right	Participants steer the steering wheel to the right	Participant press on the brake (middle paddle)

Undesired actions	Participants steer the wheel to right. Participants press the brake pedal. Non-action	Participants steers the wheel to left. Participants pressed the brake pedal. No action	Participants steers the wheel to left. Participants pressed the brake pedal. No action	Participants steers the wheel to left. Participants steers the wheel to right. Not action

Figure B1

RSME scale



Note. The nine anchor points are printed in the text.

Table B2

Factor and item description.

Factors	Description	Reference
Sleepiness	Karolinska Sleepiness Scale (KSS)	Åkerstedt & Gillberg (1990)
Trust	I trust the automation in this situation.	Adapted from Lu et al. (2022)
Mental workload	Rating Scale Mental Effort (RSME)	Adapted from Zijlstra (1993)
Trait Anxiety	I tire quickly.	Spielberger et al. (1971) as cited in Lu et al. (2022)
	I worry too much over something that really doesn't matter.	
	Some unimportant thought runs through my mind and bothers me.	
	I am a steady person.	
State Anxiety	I feel calm. I feel nervous.	Adapted from Lu et al. (2022)
	I am tense.	

Appendix C

Experimenter manual

Experimental manual

Verbal Instruction

Dear participant, thank you for participating this study on assessing the importance of trust in autonomous vehicles. The whole experiment will take approximately 90 minutes.

(ONLY TO SONA PARTICIPANTS) This study worth 2 SONA credits, you will receive your SONA credit after you have completed the study.

This study involves the use of a VR headset together with a simulator in which you might be experiencing **motion sickness** (e.g., dizziness, feeling to vomit), please immediately report to us if you experience any discomfort before, during and after the experiment.

You will have the right to withdraw this experiment at any moment without any reason, your data will also be removed. If you wish to have a copy of the informed consent, please inform us. If you have any questions up to this point, please let us know.

In this study, you will be experiencing **Level 3 autonomous vehicles**, with by simple definition the vehicle will mainly be controlled by the automation system and you as a driver are expected to take over when needed. So, in this study, you will be asked to respond to the task displayed on the screen on your right-hand side. You can respond by turning the steering wheel in the direction of left or right. You can step in the middle paddle to stop. The entire experiment contains 7 scenarios. After each scenario you will be filling in a questionnaire. Do you have any questions for now?

Participant guide

Before we began, we would like to address a few things:

- Please kindly put your phone on silence mode and place it away from your pocket, so it will not hinder you during the experiment.
- You can adjust the sitting position that best suits you by pulling the bar underneath the chair.
- Please relax and sit back during the experiment.
- Please kindly place your dominant feet on the middle paddle to stop the car.
- Please make sure you are in a sitting position + VR position where you can clearly and fully see the steering wheel and the monitor at your right-hand side in VR environment, you can also request us to adjust if these visualisations are unclear for you.
- Please note that there will be a certain level of shakiness in the vision due to technical difficulties we encounter, please don’t let in concern you.

- During the experiment you will be seeing a light grey square, it was used to perform eye tracking, please don't let it concern you.
- At the end of the experiment, the vehicle will continue to run, it may crash, or go off the road due to technical difficulties, please do not let it concern you as well.
- The program might pause or glitch due to the unity program, this will not affect the experiment, so don't let it concern you.
- Once again, if you felt uncomfortable during the experiment, please report to us immediately. We don't wish that participant to feel sick during the experiment, therefore you are free to withdraw anytime.
- You don't need to steer the steering to the max or to the hard end but make sure your action was obvious and visible to the researchers, **you can relax your arm and place it on your lap or other places, please do not place it on the steering wheel due to the setting of L3 autonomous cars.**
- We will verbally notify you whether you fail or successfully complete the task.
    - **Success ->** You have successfully take over
    - Fail -> You missed the takeover moment.
- The instruction of take-over will appear at the screen of your right hand side, please response to it as soon as you see the instruction.
- Lastly, please remember that there is no intention of assessing your driving behaviour, the aim was to assess the driving system, therefore, if you fail to perform the takeover, you don't have to worry about it (say this at the beginning of the experiment and in between, or whenever the participants seem stressed about their performance)

Notes to experimenters

- **Inform the participant whether they success or failed + input it in the data sheet**
- **Inform the participants that the aim of the experiment is not to assess their driving behaviour, but was intended to assess the system. As we encounter some participant who felt destressed and disappointed when the failed to takeover.**
- Make sure all switches are turned on before the experiment starts and are turned off when leaving
    - Computer switch
    - Simulator switches
- Make sure both of the VR sensors are in **Green**
- **Always remember to record and end record from iMotion**
- **Always to calibrate each time, after VR headset are placed back on**
- Make sure the VR headset wire is not resisting or preventing the participant to move their head freely. (Best position: the wire are placed at the left shoulder of the participants)
- Make sure the participant can reach the paddle (chair are moveable, both front/back)
- **Make sure the participant place their feet on the middle paddle and hands not on steering wheel before start**
- **Always** check the data after each trial (make sure it is recorded and copy into USB stick at the end of the day in case of data lost)

- Make sure you have inputted the data in the data sheet for chat box group
- STAY ATTENTIVE ON PARTICIPANTS ON THEIR MOTION SICKNESS, AS SOME PARTICIPANTS MIGHT FELT HESITATE TO ADDRESS IT
- IF POSSIBLE, OPEN THE DOOR OF THE SIMULATOR ROOM DURING THE BREAK FOR VENTILATION --> AVOID MOTION SICKNESS.
- **IF THE PARTICIPANTS DIDN'T ASK FOR A BREAK STILL SUGGEST A LONGER BREAK (5 MINUTES) TO AVOID MOTION SICKNESS, DURING THIS BREAK SUGGEST THE PARTICIPANT WALK AWAY FROM THE CHAIR -> REDUCE MOTION SICKNESS**
    - **EXPERIMENTER CAN CHAT WITH THE PARTICIPANTS**
    - **Offer water, not coffee or tea as sleepiness are assessed**
- if one out of 3 (or 2) experimenter say failed it will consider as a fail, if the actions was not obvious enough (so so), take it as a fail
    - Remember the participants should response **within** the time frame
        - 1 second
        - 3 seconds
        - 5 seconds
- If the program is not working or pauses, close the program and restart it. Or restart the whole PC.

<u>What to do when participants felt motion sickness</u>

- Take break + leave the room
- Take deep breath + close eyes
- Seat on the chair (not the simulator, a normal chair)
- Walk around if the participant feels okay
- Eat or drinking something
- Open door for ventilation

Appendix D
Observation sheet

Participant No.	Path	Kerem	Bart	Yumi	Notes	Final
2 (Pilot)	T		0	0		0
4/5/2023	T		1	1		1
	B (5)		0	0		0
	B		1	1		1
	B	0		0		0
	C (3)	1		1		1
	C		1	1		1
	C	1		1		1
	A (5)		1	1		1
	A		1	1		1
	A		-	-	system does not show event	NA
	D (3)		1	1		1
	D		1	1		1
	D		1	1		1
	F (1)		1	1		1
	F		1	1		1
	F		1	1		1
	E (1)		1	1		1
	E		1	1		1
	E		1	1		1

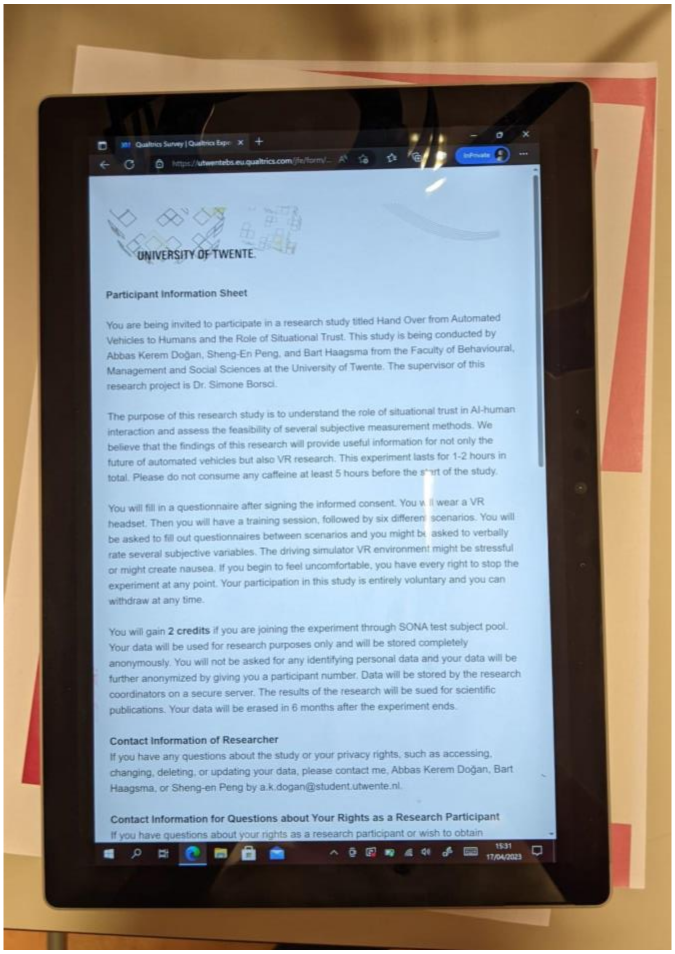
Note. The data presented are for pilot testing, this data is not used or analysed for this result. The numbers in the brackets represent the seconds provided for participants to take over. 1 indicated that the participant has successfully takeover, whereas 0 indicated failed to take over.

Appendix E
Subjective measurement questionnaire

This appendix consists of the overview of and the set-up of the survey.

Figure E1

Questionnaire displayed in Microsoft Surface Pro 9



Link to access the questionnaire

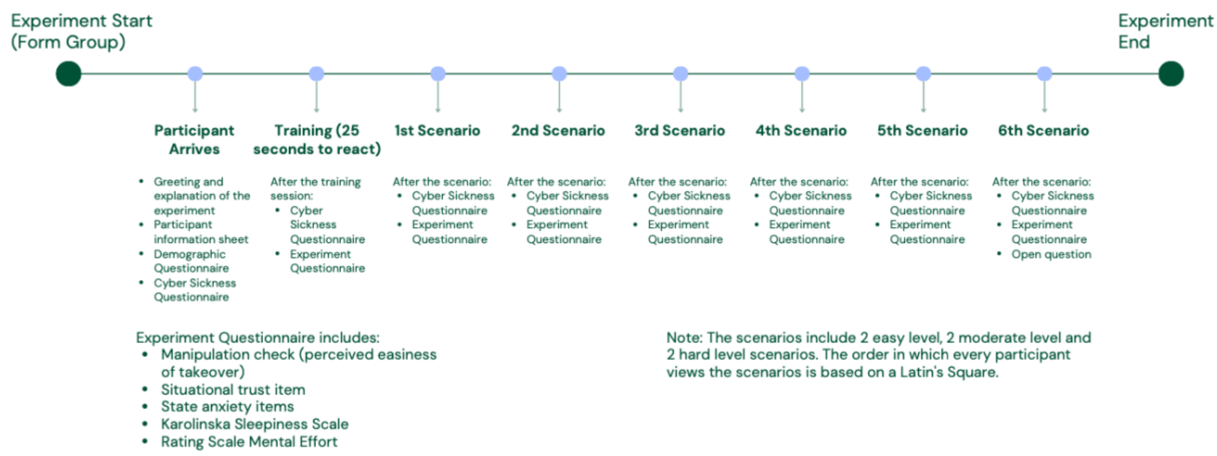
https://drive.google.com/drive/folders/1QQenAZmySfS2z_hbql5YAQ4ne41qn-L1?usp=drive_link

Appendix F

Procedure illustration

Figure F1

Overview of the experimental procedure



Appendix G

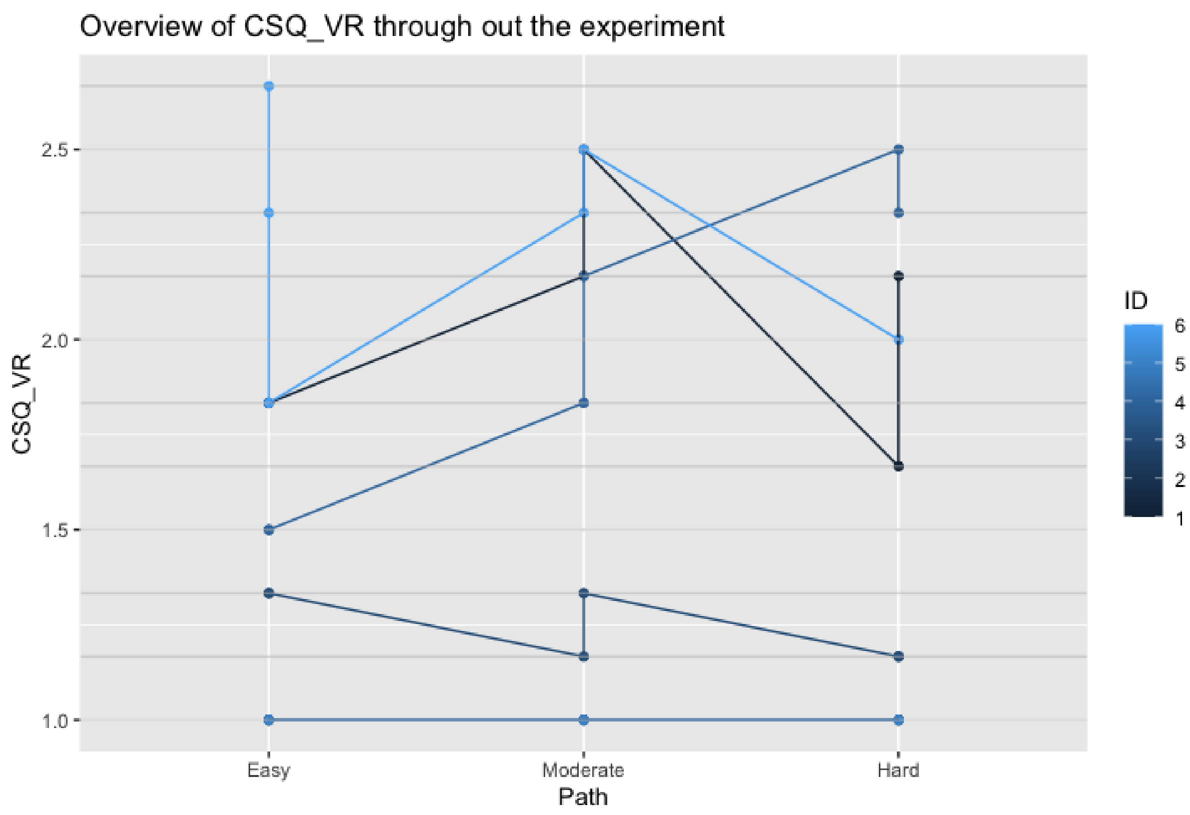
Additional data analysis

Average Motion Sickness Throughout the Experiment

Based on Figure G1, participants reported a maximum perceived motion sickness score of 2.6 on a 7-point scale. This indicates the highest level of subjective discomfort experienced during the experimental conditions. On the same 7-point scale, the lowest perceived motion sickness score reported by participants was 1, indicating the lowest level of subjective discomfort.

Figure G1

Average motion sickness throughout the experiment



Note. The scale of CSQ_VR lays between 1 (Absent feeling) to 7 (Extreme feeling).

Figure G2

Visualisation of between Level of difficulty and Situational trust

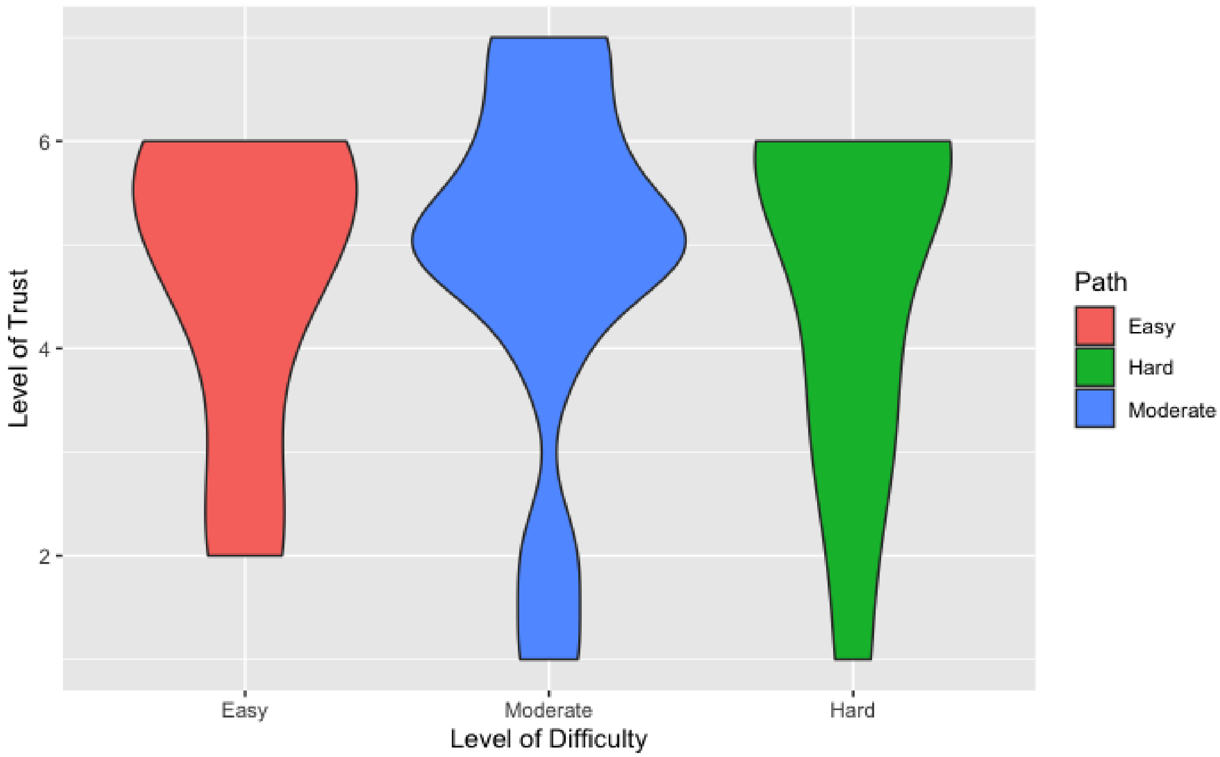


Figure G3

Visualisation of between Level of difficulty and Sleepiness

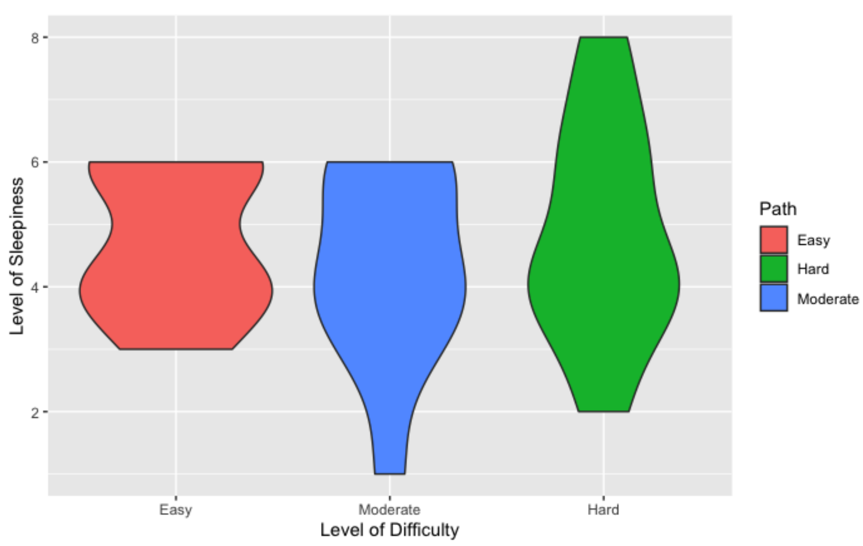


Figure G4

Visualisation of between Level of difficulty and State Anxiety

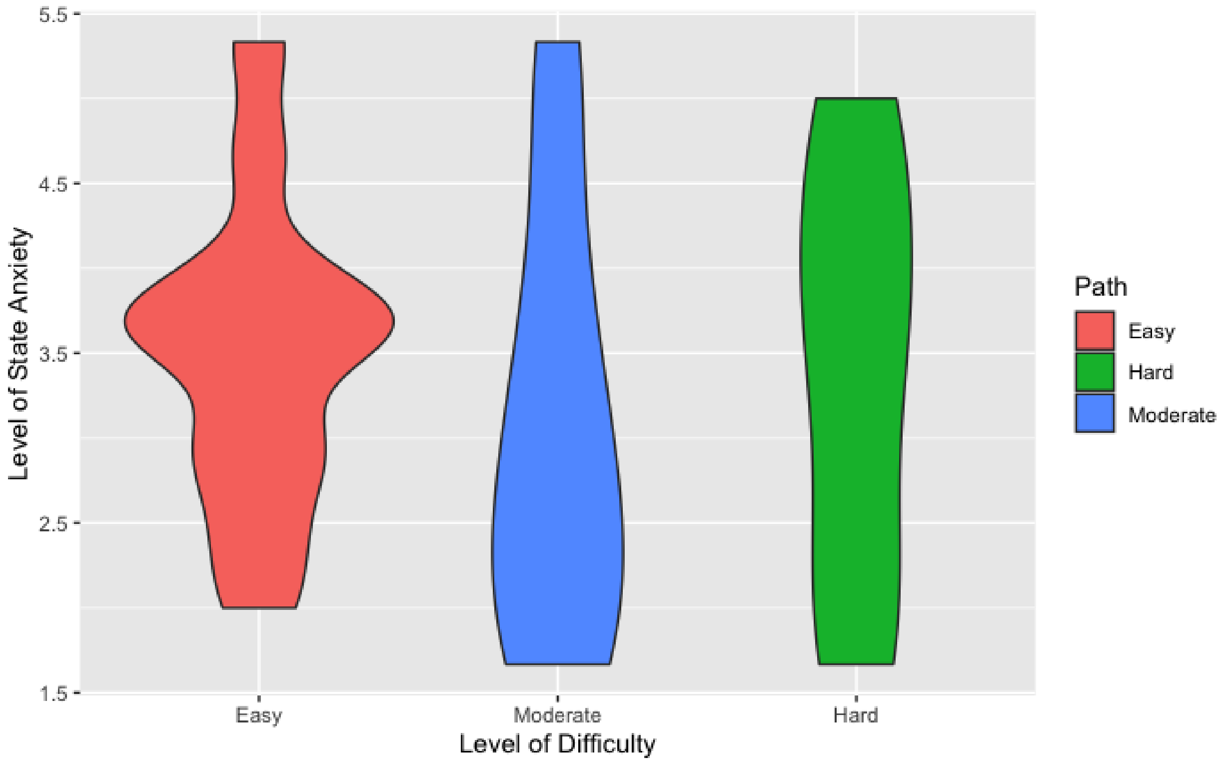
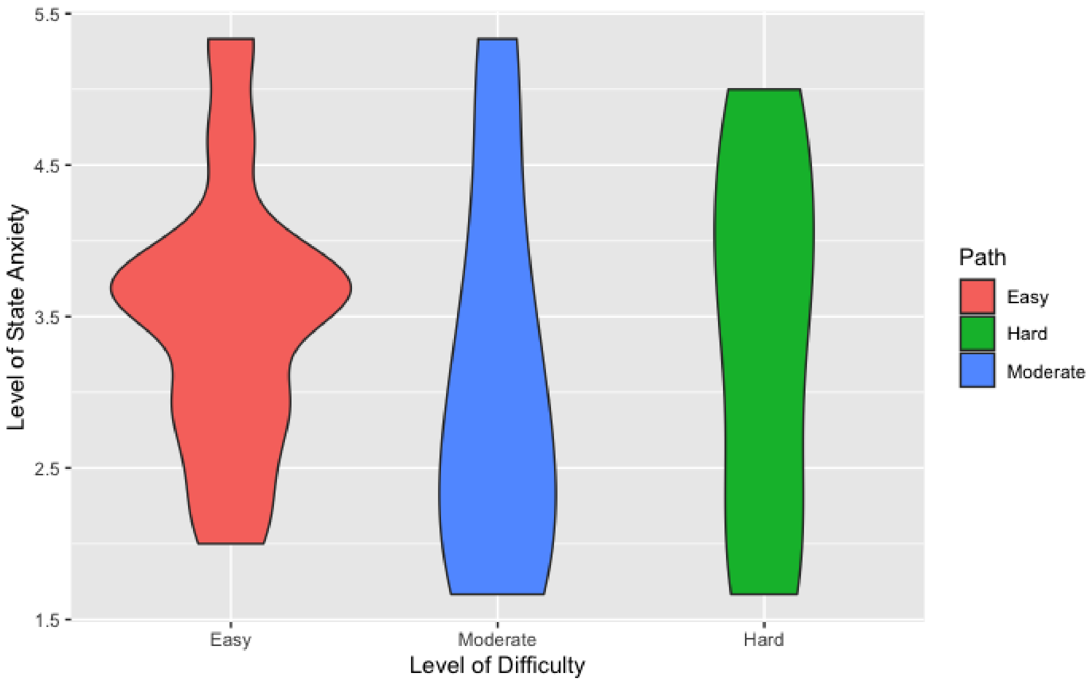


Figure G4

Visualisation of between Level of difficulty and State Anxiety



5.5
4.5
3.5
2.5
1.5
Level of State Anxiety
Easy
Moderate
Hard
Level of Difficulty
Path
Easy
Hard
Moderate

Appendix H
R code

Link to R code

https://drive.google.com/drive/folders/1QQenAZmySfS2z_hbql5YAQ4ne41qn-

L1?usp=drive_link