# A Pilot Study Exploring the Effect of Handover at Different Levels of Difficulty on Trust, Workload, Anxiety, and Sleepiness, Measured with a Conversational Agent

Bart Haagsma

Department of Psychology, University of Twente

202000384: BSc Thesis PSY

Dr. Simone Borsci

June 28, 2023

**Abstract**

The current pilot study proposed a theoretical model to investigate the relationship between Take-Over Request (TOR) difficulty and success rate, concurrently assessing the influence of trust, mental workload, state anxiety, and sleepiness. As highly autonomous vehicles become more widely available, insights into autonomous vehicle (AV) take-overs could help guide safety measures and guidelines. The main predictions of the current study were that TOR difficulty negatively affects its success, moderated by the driver's trust in the AV. Using a driving simulator and VR headset the participants completed six scenarios, each consisting of three take-over events. Each event was followed by verbal evaluations by a male researcher acting as a conversational agent, while each complete scenario was followed by a questionnaire.

Findings suggest a strong fit of the theoretical model and analyses support the collection of multiple verbal evaluations after each event rather than measuring only the final event or using a questionnaire. Results suggest that handover time and trust in the AV increases sleepiness, mental workload negatively affects TOR success, and that sleepiness correlates with mental workload and state anxiety. Contrary to existing literature however, lower handover time increased driver trust towards the AV. Given the small sample size the results should be taken with caution. Further investigations should explore alternative variable relationships.

**A Pilot Study Exploring the Effect of Handover at Different Levels of Difficulty on Trust, Workload, Anxiety, and Sleepiness, Measured with a Conversational Agent**

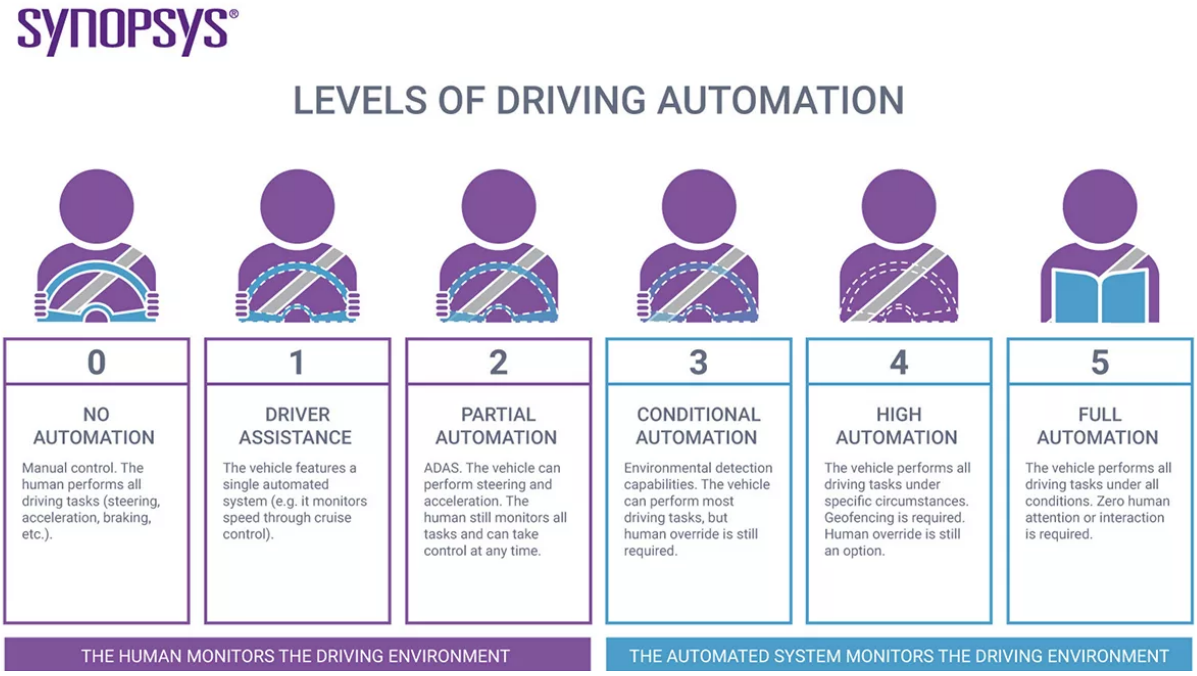
Over the past century there have been continuous interests in the creation and advancement of Automated Vehicles (AVs; Bimbraw, 2015). From the first radio-controlled car in 1926 to the Autonomous Land Vehicle (ALV) project of the 1980s, AVs have increased in complexity and capability. In more recent years, technological developments have allowed AVs to evolve into intelligent machines, able to understand their environment, adhere to traffic laws, and exchange information with other devices (Parekh et al., 2022). The adoption of AVs is predicted to have wide-ranging consequences, including impacts on future mobility, urban design, and city infrastructure (Golbabaei et al., 2020; Yigitcanlar et al., 2019). The current study aims to investigate the way humans and AVs interact in situations where the automation is unable to act appropriately, which may help guide respective safety regulations.

## Levels of Automation and Take-Over Requests

In line with the growing relevance of AVs, a necessity grew to differentiate between the different levels of automation in AVs (Hopkins & Schwanen, 2021). Whereas some vehicles imply no need for human intervention, others are limited in their autonomous capabilities. Accordingly, the Society of Automotive Engineers (SAE) has developed a framework that can be used to categorize AVs according to their level of automation (SAE, 2021). The SEA taxonomy consists of 6 levels, with level 0 connotating to no driving automation, and level 5 signifying full driving automation (see Figure 1; Synopsys, n.d.). Furthermore, each level in-between represents a significant step in either driver support or automated driving. While the three lowest levels emphasize the need for the driver to monitor the driving environment, the three highest levels rely on the autonomous system to perform this job.

**Figure 1**

*Diagram of the Different Levels of Automation*




A concept that is specifically relevant to levels 1 through 4 of the framework is the Take-Over Request (TORs; H. Clark et al., 2017; J. R. Clark et al., 2019; Eriksson & Stanton, 2017). These requests occur when an AV encounters a situation in which it is unable to respond appropriately. These situations can vary greatly in urgency, which is largely reflected in the construct of handover time, alternatively known as take-over time, response time, or lead time (Bazilinskyy et al., 2018). This term refers to the amount of time available for a driver to respond to the situation by taking over control of the vehicle, which is integral in determining the success rate of the take-over (Wan & Wu, 2018).

## Purpose of the Study

The current study is concerned with investigating the relationship between the difficulty of a TOR and its success, taking into account the trust, mental workload, sleepiness, and anxiety of the driver. Although no widespread fully automated driving systems exists, vehicles with high degrees of autonomy are expected to enter the market in the near future (European Commission & Joint Research Centre, 2019). Moreover, as the level of automation increases a respective growth in AV adoption is expected, both for commercial and industrial ends. Until widespread full automation is achieved, TORs will play an integral role in road safety. As such, the next section will elaborate on the aforementioned factors influencing the success rate of TORs, explaining the concepts themselves and how they contribute to the success overall.

## Modelling Take-Over Requests

While handover time has been linked to both the rate and quality of success, there are a lot of other factors that influence this relationship (Du et al., 2020; Wan & Wu, 2018). Although there are external variables that contribute to the outcome of a TOR, such as the level of automation and the context of situation, a large part of the possible success of a TOR is attributed to psychological factors surrounding the driver.

### *Trust in Automation*

One such factor is trust in automation, defined by Körber et al. (2018) as the willingness of a person to delegate an important task to an autonomous system, in spite of the possible repercussions. This same study shows that such trust is related to poorer take-over performance. It was found that, when transitioning from a Non-Driving Related Task (NDRT) to taking control of the vehicle, higher trust in the autonomous vehicle generally increases the time necessary to take-over, leading to more dangerous situations (Körber et al., 2018). A study by Yousfi et al. (2021), in turn, found that trust in AVs is dictated by the amount of time available for a driver to respond to a TOR, with longer lead times being associated with higher levels of trust. Moreover, trust has also been shown to influence mental workload, sleepiness, and state anxiety, each of which have their own interaction with hand-over time and success rate (Kundinger et al., 2019; Lu et al., 2022; Yousfi et al., 2021).

### *Mental Workload*

Mental workload represents how much of one's information processing capacity is occupied by neurophysiological, perceptual, and cognitive processes (Silva, 2014). Moreover, a study by Dogan et al. (2019) found that the mental workload experienced by the driver of an autonomous vehicle depends on the task they are requested to take over, with more critical situations being associated with higher mental workloads. Subsequently, higher mental workloads are linked to longer take-over times and worse take-over quality, though the extent of this effect depends on the specific source of the cognitive load (Bueno et al., 2016; Gold et al., 2016). Similarly, a study by Du et al. (2020) found that high cognitive load and short TOR lead time are associated with lower readiness and worse take-over performance. Moreover, Yousfi et al. (2021) found that the driver's level of trust also affects their mental workload, with increased trust leading to lower levels of physical, cognitive, and temporal workload.

### *State Anxiety*

Similarly, evidence also exists for a relation between trust and state anxiety of the driver (Lu et al., 2022). Whereas trait anxiety describes a person's propensity towards feeling anxious, state anxiety represents the temporary physiological, mental, and emotional effects resulting from a person's worries regarding an event or state (Spielberger & Smith, 1966). In the context of AVs, more demanding take-over situations have been linked to higher levels of anxiety (Schmidt-Daffy, 2013). Likewise, concerning the aforementioned dynamic between state anxiety and trust however, a study by Lu et al. (2022) found that the two constructs are related. Specifically, research indicates that increased levels of anxiety are predictive of lower trust in the automated driving system (Kraus et al., 2020; Lu et al., 2022; Miller et al., 2021).

### *Sleepiness*

Another construct that was found to influence the relation between TOR lead time and success rate is sleepiness. A study by Vogelpohl et al. (2019) found that driving in an AV increases fatigue experienced by the driver faster than when driving non-automated vehicles. The same study found that increased sleepiness causes drivers to react more slowly to TORs. Similar to the previous constructs, sleepiness has also been linked to trust. A study by Kundinger et al. (2019) showed that drivers with higher levels of trust in an AV exhibit more intense signs of drowsiness, consequently worsening the driver's TOR performance.

## Research Question

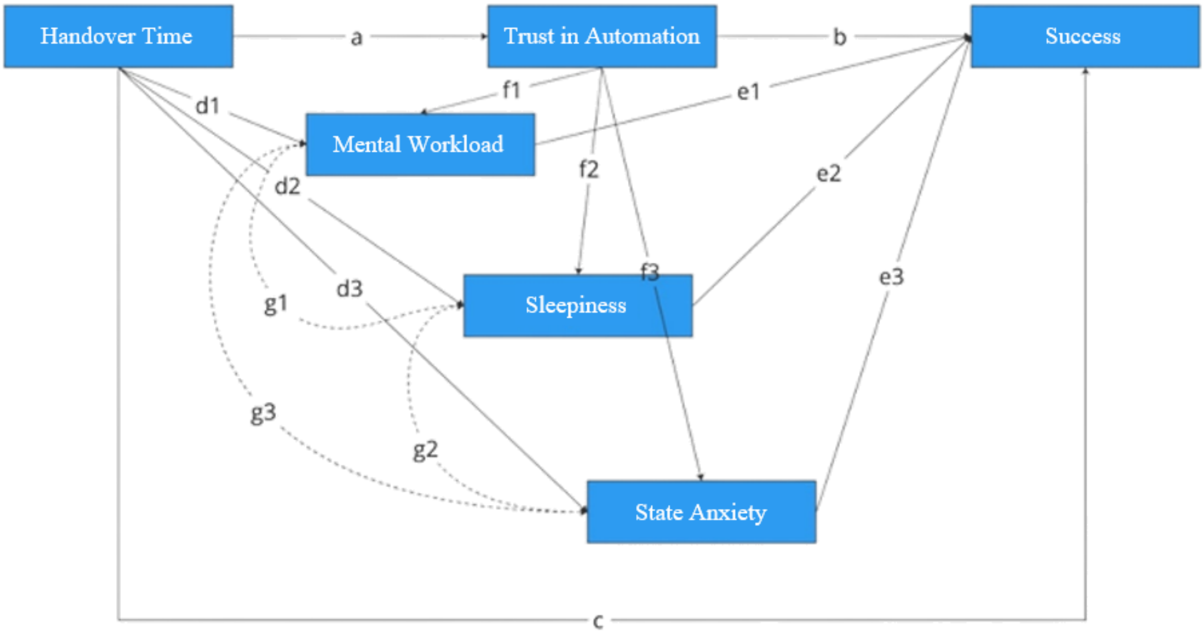
As of the current study, there is a gap in existing literature when it comes to integrating all of the previous constructs into a single model. As such, the current paper is concerned with the following research question:

How do situational trust, mental workload, state anxiety, and sleepiness affect drivers' autonomous vehicle take-over success in VR as measured through a chatbot?

Taking into account the aforementioned literature, a model of the different constructs and their relationships was created (see Figure 2). While no previous studies inform us about the possible relation between mental workload, sleepiness, and state anxiety, we will be exploring whether one exists in the context of the model. In a similar sense, since trust and anxiety have been found to correlate, it was decided to explore the possible effect of situational trust on anxiety.

**Figure 2**

*Relational Model of the Constructs Relevant to Take-Over Requests and Their Success*




Prediction 1 (a): There will be a significant positive effect between handover time and situational trust, moderating the relationship between handover time and success.

Prediction 2 (b): There will be a significant negative effect between situational trust and success, moderating the relationship between handover time and success.

Prediction 3 (c): There will be a significant negative effect between handover time and success.

Prediction 4 (d1): There will be a significant negative effect of handover time on mental workload, moderating the relationship between handover time and success.

Prediction 5 (d2): There will be a significant effect of handover time on sleepiness, moderating the relationship between handover time and success.

Prediction 6 (d3): There will be a significant negative effect of handover time on state anxiety, moderating the relationship between handover time and success.

Prediction 7 (e1): There will be a significant negative effect of mental workload on success, moderating the relationship between handover time and success.

Prediction 8 (e2): There will be a significant negative effect of sleepiness on success, moderating the relationship between handover time and success.

Prediction 9 (e3): There will be a significant effect of state anxiety on success, moderating the relationship between handover time and success.

Prediction 10 (f1): There will be a significant negative effect of situational trust on mental workload, moderating the relationship between situational trust and success.

Prediction 11 (f2): There will be a significant positive effect of situational trust on sleepiness, moderating the relationship between situational trust and success.

Prediction 12 (f3): There will be a significant negative effect of situational trust on state anxiety, moderating the relationship between situational trust and success.

Prediction 13 (g1): There will a significant negative correlation between mental workload and sleepiness.

Prediction 14 (g2): There will be a significant negative correlation between sleepiness and state anxiety.

Prediction 15 (g3): There will be a significant positive correlation between mental workload and state anxiety.

In terms of methodology, the current study aims to include a conversational agent as a means of gathering data on the subjective measures. With legislative entities trying to implement monitoring features capable of providing frequent and timely data about the driver's state, conversational agents may prove capable of performing this task (Regulation 2019/2144; Stay Aware For Everyone Act, 2021). Since driving involves comparatively less processing of auditory information compared to visual information, conversational agents may provide a good way to collect relevant data continuously without interfering much with the driving task itself (Wang et al., 2020). As such, it was decided to use a conversational agent to gather data while driving, alongside a questionnaire after the driving task is finished, to determine the best approach for collecting data about the driver.

## Methods

### *Participants*

A total of 6 participants were gathered through convenience sampling, a portion of which received credits in exchange for their participation through the University of Twente's participant pool management system for the Behavioural, Management, and Social Sciences (BMS) faculty (utwente.sona-systems.com). The current report uses a subset of participants from a larger set of individuals collected as a joint effort with two other researcher performing separate research on autonomous driving in VR.

### *Materials*

For the experiment itself, various equipment was used (see Appendix A). The environment of the driving task consisted of a driving simulator and a Vario VR-3 headset. Furthermore, a tablet was used to fill in a questionnaire at end of each trail/scenario of the experiment. Moreover, the participants were verbally provided with information and a set of instructions regarding the experiment (see Appendix B), in addition to an information sheet that was handed to the participants afterwards (see Appendix C).

A questionnaire was constructed based on the included scales to measure a number of variables over time throughout the experiment (see Appendix D). Additionally, a set of questions and statements for measuring the subjective factors verbally was created (see Appendix E).

In terms of the scales used for the subjective measures, firstly the Karolinska Sleepiness Scale (KSS) was adapted to measure sleepiness (Åkerstedt & Gillberg, 1990). It consists of a 9-point Likert-scale ranging from “Extremely alert” to “Very sleepy, great effort to keep awake, fighting sleep”. In terms of validity, the scale shows significant correlation with EEG, behavioural, and other subjective measures of sleepiness, in addition to considerable reliable (Kaida et al., 2006). Furthermore, an adapted version of the Rating Scale Mental Effort (RSME) was used to measure mental workload. It consists of a sliding scale from 0 (“Absolutely no effort”), to 35 – 57 (“Some effort to rather much effort”), to 150 (“Extreme effort”). On the whole, the scale is reliable ($r = .78$) and displays good validity through high correlation with measures of subjective effort ($r > .55$; Zijlstra, 1993). Next, an adapted form of the trust and anxiety items from Lu et al. (2022) was used to measure trait anxiety, state anxiety and situational trust. All four and three items were included for measuring trait and state anxiety respectively, while one of four total items was included to measure trust. Each of these are scored on a 7-point Likert scale ranging from “Strongly disagree” to “Strongly agree”. Regarding its psychometric properties, the items for trait ($\alpha = .90$) state anxiety ($\alpha = .95$), and trust ($\alpha = .83$) were shown to have good reliability and satisfactory aggregation validity (Lu et al., 2022). Lastly, the Cybersickness in Virtual Reality Questionnaire (CSQ-VR) was employed from Kourtesis et al. (2023) was adapted to measure cybersickness. The questionnaire consists of 6 items on nausea, vestibular, and oculomotor cybersickness, scored on a 7-point Likert scale ranging from “Absent Feeling” to “Extreme feeling”. In terms of internal consistency it has good reliability ($\alpha = .87$), while its scores also correlate strongly with those of the VRSQ ($r = .77$), supporting its convergent and construct validity (Kourtesis et al., 2023).

***Procedure***

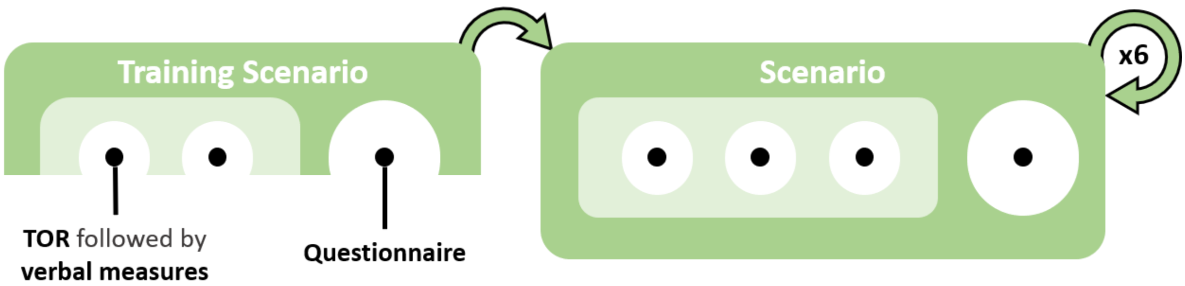
At the onset of each experiment the participant was greeted and provided with an explanation of the experiment and what their role would be, in addition to specific instructions about how to operate within the driving simulator (see Appendix B). After this initial explanation, they were also handed an information sheet with a more succinct overview of the experiment and the relevant instructions. The participant was then handed the tablet with the questionnaire, where they would fill in the informed consent, the demographic questionnaire, and a pre-emptive cybersickness questionnaire. When finished, the VR headset was mounted on the participant's head, after which the initial training scenario would ensue.

Including the training scenario, the participant would be presented with a total of seven scenarios, each of which was followed by a questionnaire (see Appendix D), conducted on the tablet after removing the headset, measuring both cybersickness and the other subjective measures (see Figure 3). Each scenario itself contained three take-over events (two in the training scenario) where the participant had to react to one of three prompts with the appropriate action in the allotted time. These actions included a right-hand turn, a stop, and a lane switch, each of which occurred once in each scenario, except for the training scenario which only included two events. The amount of time for the participants to perform these actions depended on the allotted handover time of the scenario, with the training scenario allowing 25 seconds, the easy scenarios allowing 5 seconds, the moderate scenarios allowing 3 seconds, and the hard scenarios allowing 1 second to react. Besides the training scenario, the participants completed two scenarios of each difficulty level. Following each take-over event, the participants received verbal questions on their subjective measures (see Appendix E). These questions were asked by a male researcher, simulating a conversational agent with a consistent manner of speech. To summarize, besides the training scenario, each participant completed six scenarios, each containing three take-over events (see Figure 3). Each event was followed by verbal questions regarding their situational trust, mental workload, state anxiety, and sleepiness. Lastly, at the end of each scenario, the participants would fill in a questionnaire also containing questions on these subjective measures, in addition to a cybersickness scale and a question to check for manipulation.

Between each scenario the researchers also urged the participant to notify them any possible nausea or dizziness, as well as allowing the participant to take a short break whenever they desired. At the end of the experiment the participant was thanked for their participation and led out of the experiment room.

**Figure 3**

*Overview of the Scenario Procedure*




## Data Analysis

The entirety of the data analysis was performed in R-Studio (version 4.3.0; see Appendix F). The data was first loaded into R-Studio using the "readxl" package (v1.4.2; Wickham & Bryan, 2023). After this the data was prepared for analysis, removing the training scenario data, reversing negatively phrased items, combining multi-item constructs into a single value, and creating success scores for each scenario.

Descriptive statistics were performed regarding the demographic information of the sample using the "dyplr" package (v1.1.2; Wickham et al., 2023), consisting of the mean and standard deviation of the participants' age, driving experience, trait anxiety, and cybersickness. Subsequently, a manipulation check was performed to check for differences between the different difficulty conditions. To this end the relationship between the scenario's handover time and its respective perceived difficulty was examined using a linear model, looking for statistically significant regression coefficients.

To answer our research question, we used a Structural Equation Modelling (SEM) approach, which combes elements of factor-, regression-, and path-analysis to test the predictions and validate the initially constructed model. More specifically, a model selection was employed, intended as an exploratory and comparative approach aiming to identify the best model to fit the data (Hoyle, 2023). To this end, three models were created. The first model used the average scores of all the verbal measures taken in each scenario. This model encompasses all the participants' changes throughout the events, under the assumption that each event is equally representative. The second model used the verbal measures of only the final event of each scenario. Compared to the first model, the second model only takes into account data from the participants after they have been driving for an extended time. Lastly, the third model used the scores retrieved from the questionnaire at the end of each scenario. Contrary to the previous two models, model three is based on data from a different modality, employing data from a questionnaire instead of a conversational agent.

The models were compared using five different parameters: the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA), the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Each of these fit indices assess how well the observed data fits the proposed model, though each using a different approach. While the RMSEA assesses how well the model fits the data in terms of an absolute score, the CFI and TLI evaluate the fit of the model by comparing it to alternative models. The RMSEA represents the average discrepancy between the model and the observed data with lower RMSEA values indicating better fit and values below 0.05 being considered indicative of good fit. The CFI and TLI on the other hand, evaluate how well the model reproduces the observed data patterns, with values closer to 1 indicating a better fit. Finally, the AIC and the BIC are known as parsimony fit indices. These parameters take into account the complexity of the model and penalize models which are overly complex, with lower values indicating better fit.

After selecting the best-fitting model, bootstrapping was used to obtain more robust estimates of the model parameters and to compensate for the small sample size. Bootstrapping enables robust analysis without strict assumptions about data distribution. It does this by generating multiple resampled datasets that mimic the properties of the original data, overcoming limitations associated with small samples in traditional SEM methods. Both the regular – for the purpose of model selection - and bootstrapped SEM models were ran and examined with the "lavaan" package (v0.6-15; Rosseel, 2012). In terms of parameters, the bootstrapped versions of both models specifically using a sample size of 5000 and a confidence interval of 95%.

## Results

Participant characteristics can be found in Table 1. The sample consisted of 6 individuals with an average age of 23 ($SD = 2.19$) and an equal distribution of male and female participants.

**Table 1**

*Participant Characteristics*

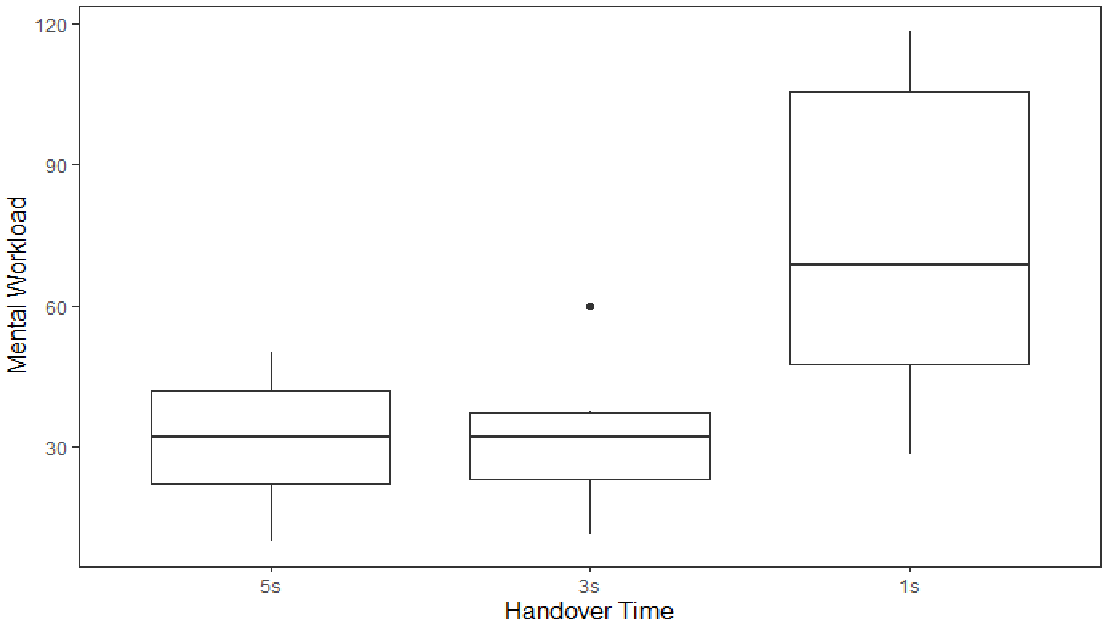
| Variable | Mean | SD |
|---|---|---|
| Driving experience | 4.33 | 2.50 |
| Trait anxiety | 3.33 | 0.44 |
| Cybersickness | 1.33 | 0.59 |

*Note.* Driving experience refers to the number of years the participant has had their driver's licence. Trait anxiety and cybersickness are both based on a 7-point Likert scale.

A manipulation check was conducted using linear regression analysis to examine the influence of the experimental manipulation condition, the handover time of the path on perceived difficulty. The results of the linear regression revealed a significant effect of the manipulation condition on perceived difficulty, $F(2, 33) = 6.43, p = .004$. Specifically, a handover time of 1 second was perceived as significantly less easy, $\beta = -1.67, SE = 0.57, t = -2.95, p = .006$, while a handover time of 3 seconds did not show a significant effect compared to a handover time of 5 seconds, $\beta = 0.17, SE = 0.57, t = 0.30, p = .770$. This is most clear when mapped against mental workload (see Figure 4). The overall model accounted for 28.03% of the variance in perceived difficulty ($R^2 = 0.28$) indicating a moderate effect size.

**Figure 4**

*Distributions of Mental Workload Across Handover Time*




These findings provide support for the successful manipulation of scenario difficulty across different conditions, demonstrating that the manipulation had the intended impact on participants' perception of task difficulty.

To examine the optimal fit for the data, three alternative structural equation models were compared. Model 1 and 2 used the data gathered through the conversational agent, model 1 using the average verbal measures, while model 2 used the verbal measures of each scenario's final event. Lastly, model 3 used the data collected from each scenario's questionnaire. See Table 2 for an overview of the indices used to compare the models.

**Table 2**

*Overview of the Fit Indices of the Models*

| Index | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| CFI | 1.00 | 1.00 | 1.00 |
| TLI | 1.00 | 1.00 | 1.00 |
| RMSEA | 0.00 | 0.00 | 0.00 |
| AIC | 155.84 | 452.57 | 444.41 |
| BIC | 173.65 | 484.24 | 476.08 |

Though no differences were found between the models' CFI and TLI scores, each indicate strong alignment between the proposed model and the observed data. Similarly, the RSMEA values of the models were all the same, each suggesting a perfect fit between the proposed model and the observed. However, differences were found between the models' AIC and BIC scores. While the model 2 and 3 performed similarly, model 1 exhibited a superior fit to the data as evidenced by its lower AIC and BIC values. As such, the best-fitting model was determined to be model 1, using the average verbal measures collected by the conversational agent throughout the entirety of each scenario.

Model 1 was assessed using a bootstrapped SEM approach. However, it should be noted that the model is overfit and suggests saturation as indicated by the perfect CFI, TLI and RMSEA criteria. Overfitting occurs when a model captures random noise or idiosyncrasies in the data, resulting in excessive complexity that may not generalize well to new data. Saturated models perfectly reproduce the observed data, fitting it with no residual discrepancies. Given the combination of a saturated model, perfect fit indices, and a small sample size, caution must be exercised when interpreting the results due to overfitting.

The results of the analysis of the selected model and the answer to the predictions are reported in Table 3.

Results show that, in terms of the direct effects, only the effect of difficulty on trust was significant. However, contrary to the expectations, handover time negatively affected the level of trust. In terms of moderation, handover time showed a significant positive effect on sleepiness. Regarding the indirect effects, there was a significant relationship between trust and sleepiness, with trust increasing the level of sleepiness. In terms of predictors of success, the mental workload was the only subjective measure to significantly influence success, with higher workloads leading to lower rates of success. Lastly, regarding correlations, the level of sleepiness was found to correlate to both the mental workload and state anxiety.

As such predictions 5, 7, 11, 13, and 15 are confirmed, while prediction 1 is rejected, though the limited generalizability of the model should be kept in mind when interpreting these results. The remainder of the predictions cannot conclusively be confirmed based on the current data.

**Table 3**

*Regression Effects and Respective Predictions of Model 1 (Bootstrapped)*

| Prediction | Relation | Model Label | Estimate | SE | z-value | p-value | Prediction Outcome |
|---|---|---|---|---|---|---|---|
| **1** | **Handover Time → Trust** | **a** | **-0.62** | **0.18** | **3.45** | **.001** | **Rejected** |
| 2 | Trust → Success | b | -0.27 | 0.20 | -1.39 | .166 | |
| 3 | Handover Time → Success | c | -0.15 | 0.15 | 0.97 | .334 | |
| 4 | Handover Time → MWL | d1 | -0.14 | 0.33 | 0.42 | .675 | |
| **5** | **Handover Time → Sleepiness** | **d2** | **0.70** | **0.18** | **-3.84** | **<.001** | **Confirmed** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 6 | Handover Time → State Anxiety | d3 | -0.33 | 0.29 | 1.11 | .267 | |
| **7** | **MWL → Success** | **e1** | **-0.57** | **0.13** | **-4.59** | **<.001** | **Confirmed** |
| 8 | Sleepiness → Success | e2 | 0.22 | 0.20 | 1.09 | .274 | |
| 9 | State Anxiety → Success | e3 | -0.09 | 0.12 | -0.75 | .454 | |
| 10 | Trust → MWL | f1 | -0.42 | 0.36 | -1.15 | .250 | |
| **11** | **Trust → Sleepiness** | **f2** | **0.73** | **0.22** | **3.35** | **.001** | **Confirmed** |
| 12 | Trust → State Anxiety | f3 | -0.43 | 0.34 | -1.26 | .207 | |
| **13** | **Sleepiness ↔ MWL** | **g1** | **0.30** | **0.11** | **2.87** | **.004** | **Confirmed** |
| 14 | State Anxiety ↔ MWL | g2 | -0.02 | 0.15 | -0.11 | .910 | |
| **15** | **State Anxiety ↔ Sleepiness** | **g3** | **-0.29** | **0.13** | **-2.25** | **.024** | **Confirmed** |

*Note.* MWL refers to mental workload

## Discussion

The current study sought to investigate the relationship between the difficulty of a TOR and its possible success, taking into account the trust, mental workload, sleepiness, and anxiety of the driver. Moreover, a conversational agent was used alongside a questionnaire to collect data about these constructs to determine the best approach for doing so.

The results of the current study show evidence in support of five out of the fifteen predicted relationships in the proposed model. In line with existing literature, it was found that higher mental workloads led to longer take-over times and lower rates of take-over success (Bueno et al., 2016; Du et al., 2020; Gold et al., 2016). Moreover, similar to the findings of Kundinger et al. (2019), the current study found that higher levels of trust in the AV resulted in more experienced sleepiness.

To our knowledge this study was the first one to attempt to assess the effect of TOR difficulty on sleepiness in AVs. In this regard, our findings align with the indications coming from Drummond et al. (2004) which suggested that an increase in task difficulty results in a more wakeful state. Similarly, the current study is the first to investigate the correlation of mental workload, sleepiness, and state anxiety in TORs. To this end, sleepiness was found to correlate with both mental workload and state anxiety, though no direct link between mental workload and state anxiety was found.

Lastly, it was expected that TOR difficulty would negatively affect the driver's trust in the autonomous vehicle, as suggested by the results of Yousfi et al. (2021). Contrary to this, our findings indicate that the difficulty increases the trust of the driver in the AV. Although this may be the result of methodological differences, it is possible that the relationship between TOR difficulty and driver trust may be more complex than assumed by the current model. While complexities such as additional moderating variables or non-linear patterns could explain the discrepancy, further research is needed to clarify the relationship between TOR difficulty and driver trust.

Overall, the proposed model should be tested with more participants and a different operationalization for the level of difficulty. Currently, the difference between two of the three level of difficulties is not significant and two of the three main direct effects were not significant. This prevented us from discussing the rejected predictions, as the present pilot was not powerful enough nor was the design optimal to provide a meaningful analysis. Overall, the proposed model seems powerful and valid, but present data should be treated with caution.

In any case, the current analysis suggests that when the model is fed with data collected through repeated verbalisations via a conversational agent, the fit is better than when data are collected less frequently or through a questionnaire. This suggests that conversational agents could be a viable means of collecting frequent and timely data about the state of the driver. In turn, this may make conversational agents a good solution for driver monitoring in AVs as urged by legislative entities (Regulation 2019/2144; Stay Aware For Everyone Act, 2021). That said, as of yet it is still unclear if and how conversing with the agent may affect the driving task or TORs, showing a need for further research.

**Limitations and Future Recommendations**

Turning to limitations of the current study, there are various issues that may have negatively impacted the validity and reliability of the results. To start, the sampling validity is lacking in both quality and quantity. The current sample, consisting of six young adults, fails to meet the recommended size for investigating the proposed model, which would be 10 participants per examined relationship. Moreover, while amongst the group most likely to adopt autonomous vehicles, young adults comprise only a portion of the population expected to use AVs in the future (Lee et al., 2017; Rovira et al., 2019).

Regarding methodological limitations of the current study, technical issues with the simulation led to TOR success being measured through observation, as opposed to determining success computationally in the simulation itself. Additionally, there were also a number of problems in the fidelity of the simulation, namely limited visual quality, shakiness throughout the driving experience, a visual square in the centre of the participant's field of view, the car crashing at the end of each path, random freezing of the simulation, and a lack of side mirror reflections. Moreover, the current study only collected qualitative data and subjective ratings. In future research, it is recommended to also include objective measures to increase objectivity, improve reliability, and aid in cross-validating the results. Lastly, in terms of the procedure, since participants were informed about their success or failure in the TOR event before being queried about their subjective measures, it is possible that their responses were influenced by their knowledge of the outcome of the task.

Regarding the data analysis, constraints in time and resources limited the scope of the current study in terms of exploring different models. The inclusion of additional constructs such as situational awareness, which has been shown to affect reaction speed, may have revealed more complex relationships or lead to improved explanatory and/or predictive ability (H. Clark et al., 2017; J. R. Clark et al., 2019). Likewise, additional exploration of model configurations may have led to more representative models or new found construct interactions altogether.

As for recommendations of future research, it is worth reinvestigating the proposed model as limitations in the current study prevented it from discovering conclusive results. Besides that, it may be valuable to investigate how additional constructs such as driving experience and situational awareness further influence the relationship between TOR difficulty and success. Moreover, additional research into the use of conversational agents for AV driver monitoring may lead to recommendations regarding the tone of voice, use of language, or other factors possibly relevant in collecting driver data. In terms of research into methodologies for studies investigating the interaction of people with autonomous vehicles, it could be worth looking into how factors like simulation fidelity and VR reflect on real-world settings.

**Conclusion**

The current pilot study sought to investigate the relationship between the difficulty of a TOR due to unexpected events and people’s success in taking over, by also considering the role of situational trust, mental workload, sleepiness, and state anxiety of the driver. Based on existing literature a model was constructed, which was later tested and analysed through a SEM. In line with literature, the difficulty of a TOR was found to decrease the driver’s sleepiness, while their trust in the AV increased their sleepiness. Moreover, a higher mental workload was found to lead to lower rates of TOR success. Additionally, sleepiness was discovered to correlate with both mental workload and state anxiety. Contrary to existing literature however, the results showed that higher take-over difficulty increased the trust of the driver towards the AV. Lastly, of the included measurement modalities, the conversational agent that collected subjective measures the most frequently throughout each scenario resulted in data with the best fit to the proposed model. While limitations in the current study restrict its generalizability, the model that we tested seems a good theoretical basis for future investigations that aim to concurrently test multiple interconnected (e.g., trust, sleepiness, etc.) aspects during unexpected tasks of taking over from an autonomous vehicle.

**References**

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

Bazilinskyy, P., Petermeijer, S. M., Petrovych, V., Dodou, D., & de Winter, J. C. F. (2018). Take-over requests in highly automated driving: A crowdsourcing survey on auditory, vibrotactile, and visual displays. *Transportation Research Part F: Traffic Psychology and Behaviour*, *56*, 82–98. https://doi.org/10.1016/j.trf.2018.04.001

Bimbraw, K. (2015). Autonomous cars: Past, present and future a review of the developments in the last century, the present scenario and the expected future of autonomous vehicle technology. *2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, *01*, 191–198.

Bueno, M., Dogan, E., Hadj Selem, F., Monacelli, E., Boverie, S., & Guillaume, A. (2016). How different mental workload levels affect the take-over control after automated driving. *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2040–2045. https://doi.org/10.1109/ITSC.2016.7795886

Clark, H., McLaughlin, A. C., & Feng, J. (2017). Situational Awareness and Time to Takeover: Exploring an Alternative Method to Measure Engagement with High-Level Automation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 1452–1456. https://doi.org/10.1177/1541931213601848

Clark, J. R., Stanton, N. A., & Revell, K. M. A. (2019). Conditionally and highly automated vehicle handover: A study exploring vocal communication between two drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *65*, 699–715. https://doi.org/10.1016/j.trf.2018.06.008

Dogan, E., Honnêt, V., Masfrand, S., & Guillaume, A. (2019). Effects of non-driving-related tasks on takeover performance in different takeover situations in conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *62*, 494–504. https://doi.org/10.1016/j.trf.2019.02.010

Drummond, S., Brown, G., Salamat, J., & Gillin, J. (2004). Increasing Task Difficulty Facilitates the Cerebral Compensatory Response to Total Sleep Deprivation. *Sleep*, *27*, 445–451. https://doi.org/10.1093/sleep/27.3.445

Du, N., Kim, J., Zhou, F., Pulver, E., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Evaluating Effects of Cognitive Load, Takeover Request Lead Time, and Traffic Density on Drivers' Takeover Performance in Conditionally Automated Driving. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 66–73. https://doi.org/10.1145/3409120.3410666

Eriksson, A., & Stanton, N. A. (2017). Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control. *Human Factors*, *59*(4), 689–705. https://doi.org/10.1177/0018720816685832

European Commission & Joint Research Centre. (2019). *The future of road transport: Implications of automated, connected, low-carbon and shared mobility*. Publications Office. https://doi.org/10.2760/668964

Golbabaei, F., Yigitcanlar, T., Paz, A., & Bunker, J. (2020). Individual Predictors of Autonomous Vehicle Public Acceptance and Intention to Use: A Systematic Review of the Literature. *Journal of Open Innovation: Technology, Market, and Complexity*, *6*(4), Article 4. https://doi.org/10.3390/joitmc6040106

Gold, C., Körber, M., Lechner, D., & Bengler, K. (2016). Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density.

*Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(4), 642–652. https://doi.org/10.1177/0018720816634226

Hopkins, D., & Schwanen, T. (2021). Talking about automated vehicles: What do levels of automation do? *Technology in Society*, *64*, 101488. https://doi.org/10.1016/j.techsoc.2020.101488

Hoyle, R. H. (Ed.). (2023). *Handbook of structural equation modeling* (Second edition). The Guilford Press.

Kaida, K., Takahashi, M., Akerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *117*(7), 1574–1581. https://doi.org/10.1016/j.clinph.2006.03.011

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*, *2*(1), Article 1. https://doi.org/10.3390/virtualworlds2010002

Kraus, J., Scholz, D., Messner, E.-M., Messner, M., & Baumann, M. (2020). Scared to Trust? – Predicting Trust in Highly Automated Driving by Depressiveness, Negative Self-Evaluations and State Anxiety. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02917

Kundinger, T., Wintersberger, P., & Riener, A. (2019). (Over)Trust in Automated Driving: The Sleeping Pill of Tomorrow? *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3290607.3312869

Lee, C., Ward, C., Raue, M., D'Ambrosio, L., & Coughlin, J. F. (2017). *Age Differences in Acceptance of Self-driving Cars: A Survey of Perceptions and Attitudes* (J. Zhou & G. Salvendy, Eds.; Vol. 10297, pp. 3–13). Springer International Publishing. https://doi.org/10.1007/978-3-319-58530-7_1

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

Miller, L., Kraus, J., Babel, F., & Baumann, M. (2021). More Than a Feeling—Interrelation of Trust Layers in Human-Robot Interaction and the Role of User Dispositions and State Anxiety. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.592711

Parekh, D., Poddar, N., Rajpurkar, A., Chahal, M., Kumar, N., Joshi, G. P., & Cho, W. (2022). A Review on Autonomous Vehicles: Progress, Methods and Challenges. *Electronics*, *11*(14), Article 14. https://doi.org/10.3390/electronics11142162

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rovira, E., McLaughlin, A. C., Pak, R., & High, L. (2019). Looking for Age Differences in Self-Driving Vehicles: Examining the Effects of Automation Reliability, Driving

Risk, and Physical Impairment on Trust. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00800

SAE. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. SAE International. https://doi.org/10.4271/J3016_202104

Schmidt-Daffy, M. (2013). Fear and anxiety while driving: Differential impact of task demands, speed and motivation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *16*, 14–28. https://doi.org/10.1016/j.trf.2012.07.002

Silva, F. P. da. (2014). Mental Workload, Task Demand and Driving Performance: What Relation? *Procedia - Social and Behavioral Sciences*, *162*, 310–319. https://doi.org/10.1016/j.sbspro.2014.12.212

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

Synopsys. (n.d.). *The 6 Levels of Vehicle Autonomy Explained | Synopsys Automotive*. Retrieved 6 June 2023, from https://www.synopsys.com/automotive/autonomous-driving-levels.html

Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis & Prevention*, *126*, 70–84. https://doi.org/10.1016/j.aap.2018.03.013

Wan, J., & Wu, C. (2018). The Effects of Lead Time of Take-Over Request and Nondriving Tasks on Taking-Over Control of Automated Vehicles. *IEEE Transactions on Human-Machine Systems*, *48*(6), 582–591. https://doi.org/10.1109/THMS.2018.2844251

Wang, M., Liao, Y., Lyckvi, S. L., & Chen, F. (2020). How drivers respond to visual vs. Auditory information in advisory traffic information systems. *Behaviour & Information Technology*, *39*(12), 1308–1319. https://doi.org/10.1080/0144929X.2019.1667439

Wickham, H., & Bryan, J. (2023). *readxl: Read Excel Files*. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr

Yigitcanlar, T., Wilson, M., & Kamruzzaman, M. (2019). Disruptive Impacts of Automated Driving Systems on the Built Environment and Land Use: An Urban Planner's Perspective. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(2), Article 2. https://doi.org/10.3390/joitmc5020024

Yousfi, E., Malin, S., Halit, L., Roger, S., & Dogan, E. (2021). Driver experience and safety during manual intervention in a simulated automated vehicle: Influence of longer time margin allowed by connectivity. *Proceedings of the 32nd European Conference on Cognitive Ergonomics*, 1–7. https://doi.org/10.1145/3452853.3452888

Zijlstra, F. R. H. (1993). *Efficiency in work behaviour: A design approach for modern tools*. https://repository.tudelft.nl/islandora/object/uuid%3Ad97a028b-c3dc-4930-b2ab-a7877993a17f

## Appendix A

**Figure 1**

*Driving Simulator Entire Rig*

**Figure 2**

*Driving Simulator Steering Wheel*

**Figure 3**

*Driving Simulator Pedals*

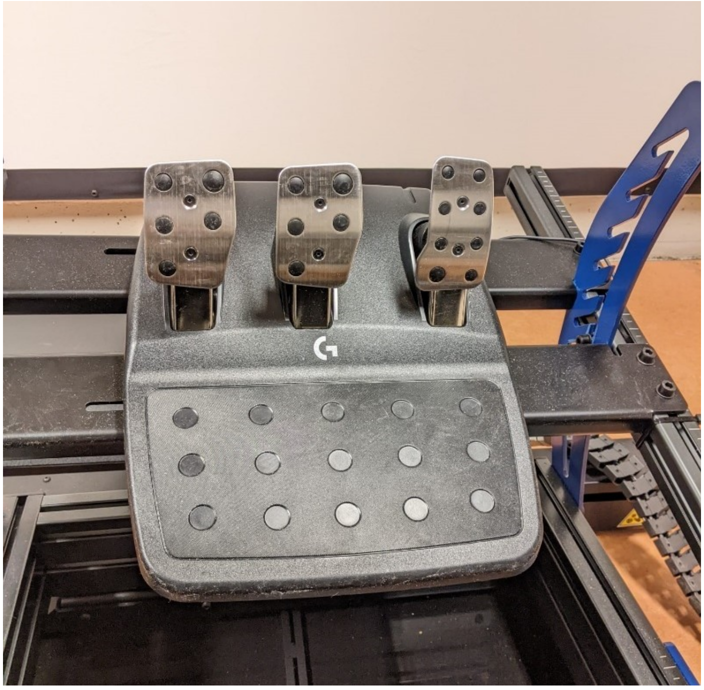
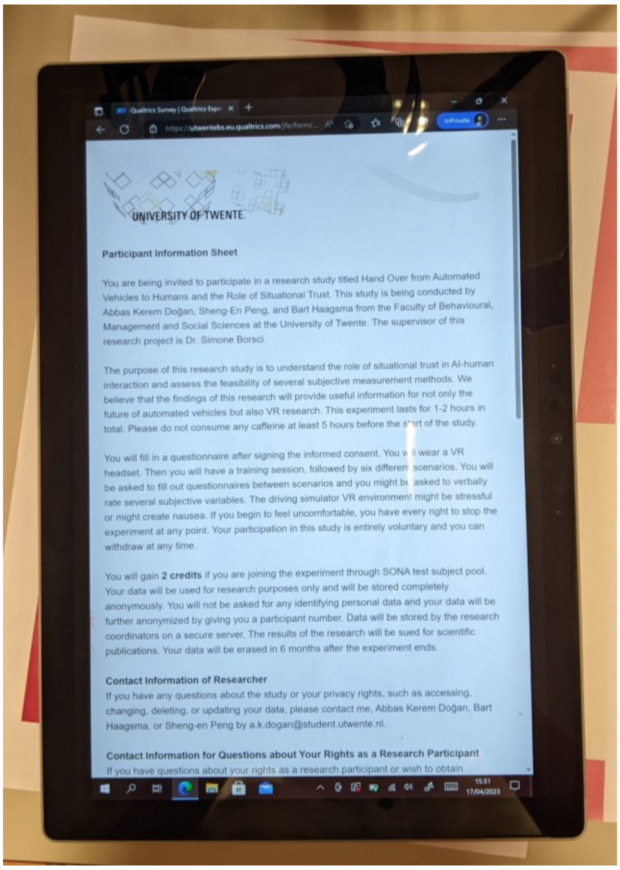


**Figure 4**

*Vario VR-3 Headset*

**Figure 5**

*Tablet Used for the Post-Trail Questionnaires*



# Appendix B

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.

(Exclusive to SONA participants) This study is 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 begin, 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 are have successfully taken over
    - *Fail* -> You missed the takeover moment
- The instruction to "take over" will appear at the screen of your right-hand side, please respond to it as soon as you see the instruction.
- Lastly, please remember that there is no intention of assessing your driving behaviour, the aim is 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)

# Appendix C

## INSTRUCTIONS

Society of Automative Engineers (SAE) defines Level 3 automation as conditional autonomy. The vehicle can operate independently. Steering functions, braking and acceleration are automated but the driver must be prepared to intervene. As an SAE Level 3 feature, the autonomous vehicle expects the fallback-ready user seated in the driver's seat to resume driving when requested to do so.

In this experiment, you will be asked to takeover from the automation during random points. Please do not touch the steering wheel or the pedals if you are not instructed to do so. The instructions are:

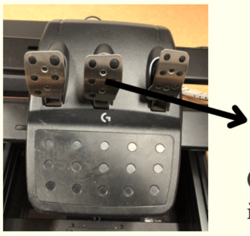
Steer the wheel to the right to move to the right lane

Steer the wheel to the right to take the right exit

Steer the wheel to the left to move to the left lane

STOP

Press on the break to stop the car



**Break**

(To stop the car when instructed to do so)

**Steering wheel**

(To exit or change lane when instructed to do so)

- If you experience nausea, dizziness, disorientation, postural instability, visually induced fatigue, and/or visually induced discomfort at any point during the experiment please inform the researchers.

# Appendix D



## Participant Information Sheet

## Participant Information Sheet

You are being invited to participate in a research study titled Hand Over from Automated Vehicles to Humans and the Role of Situational Trust. This study is being conducted by Abbas Kerem Doğan, Sheng-En Peng, and Bart Haagsma from the Faculty of Behavioural, Management and Social Sciences at the University of Twente. The supervisor of this research project is Dr. Simone Borsci.

The purpose of this research study is to understand the role of situational trust in AI-human interaction and assess the feasibility of several subjective measurement methods. We believe that the findings of this research will provide useful information for not only the future of automated vehicles but also VR research. This experiment lasts for 1-2 hours in total. Please do not consume any caffeine at least 5 hours before the start of the study.

You will fill in a questionnaire after signing the informed consent. You will wear a VR headset. Then you will have a training session, followed by six different scenarios. You will be asked to fill out questionnaires between scenarios and you might be asked to verbally rate several subjective variables. The driving simulator VR environment might be stressful or might create nausea. If you begin to feel uncomfortable, you have every right to stop the experiment at any point. Your participation in this study is entirely voluntary and you can withdraw at any time.

You will gain **2 credits** if you are joining the experiment through SONA test subject pool. Your data will be used for research purposes only and will be stored completely anonymously. You will not be asked for any identifying personal data and your data will be further anonymized by giving you a participant number. Data will be stored by the research coordinators on a secure server. The results of the research will be sued for scientific publications. Your data will be erased in 6 months after the experiment ends.

**Contact Information of Researcher**
If you have any questions about the study or your privacy rights, such as accessing, changing, deleting, or updating your data, please contact me, Abbas Kerem Doğan, Bart Haagsma, or Sheng-en Peng by a.k.dogan@student.utwente.nl.

**Contact Information for Questions about Your Rights as a Research Participant**
If you have questions about your rights as a research participant or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s). You can also contact for additional information the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management, and Social Sciences at the University of Twente by ethicscommittee-hss@utwente.nl.

| | Yes | No |
|---|---|---|
| I have read and understood the study information about the study, or it has been read to me. | ☐ | ☐ |
| I was informed that I may ask questions about the study, or I asked questions by email about the study and my questions have been answered to my satisfaction. | ☐ | ☐ |
| I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason. | ☐ | ☐ |
| I understand that biophysical data are going to be collected in this study. | ☐ | ☐ |
| I understand that the information I provide will be used for educational purposes, reports, and potential publications. | ☐ | ☐ |
| I understand that personal information collected about me that | ☐ | ☐ |

| | Yes | No |
|---|---|---|
| can identify me, such as my name, etc, will not be shared beyond the study team. | | |
| I give permission for the collected data that I provide to be archived in the Qualtrics system and in a secure hard drive so it can be used for future research and learning. | ☐ | ☐ |

**Demographic Questionnaire -1**

Participant Number

Age

What is your sex as assigned at birth?

- ○ Male
- ○ Female

What is your current current gender identity? (Please check all that apply)

*Information associated with this question is not going to be used or shared for the research

**This question is optional and could be skipped

*** This question was developed in tune with: Broussard, K. A., Warner, R. H., & Pope, A. R. (2018). Too many boxes, or not enough?

Preferences for how we ask about gender in cisgender, LGB, and gender-diverse samples. Sex Roles, 78(9), 606-624.

- ☐ Man
- ☐ Woman
- ☐ Female-to-Male (FTM)/Transgender Male/Trans Man
- ☐ Male-to-Female (MTF)/Transgender Female/Trans Woman
- ☐ Genderqueer, neither exclusively male or female

☐ Additional Gender Category/(or Other), please specify

☐ Decline to answer

**Demographic Questionnaire -2**

Please indicate for each statement whether it applies to you by checking the box.

☐ I have not consumed any alcohol in the past 24 hours

☐ I have not consumed any caffeine in the past 5 hours

☐ I am not pregnant, nor I am at risk of pregnancy

☐ I do not have colorblindness

☐ I do not have any eye condition that might affect the interaction with the virtual system like for instance presbyopia, amblyopia

☐ I have a driving license (Please specify the amount of years of driving experience)

☐ I have had no prior issues with Virtual Reality. (If so, please specify:)

Please indicate for each statement whether it applies to you by checking the box.

☐ I have not consumed any alcohol in the past 24 hours

☐ I have not consumed any caffeine in the past 5 hours

☐ I do not have colorblindness

☐ I do not have nay eye condition that might affect the interaction with the virtual system like for instance presbyopia, amblyopia

☐ I have a driving license (Please specify the amount of years of driving experience)

☐ I have had no prior issues with Virtual Reality. (If so, please specify:)

Please read the below statements carefully and choose the option that you think fits you the best.

| | Strongly disagree | Disagree | Somewhat disagree | Neither agree nor disagree | Somewhat agree | Agree | Strongly agree |
|---|---|---|---|---|---|---|---|
| I get tired quickly. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| 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. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

**CSQ-VR**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light-headedness or spinning feeling)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced fatigue (e.g., feeling | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| of tiredness or sleepiness)? | | | | | | | |
| Do you experience a visually induced discomfort (e.g., eyestrain, blurred vision, or headache)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please write below any additional comments relevant to the question above:

**Wait -Training**

Now, you will receive further instructions from the researcher. **Please inform them that you have filled in the questionnaire and do not leave this page unless instructed to do so.**

Afterwards, when the researchers tells you, you can come back to this questionnaire and fill out the remaining questions.

**CSQ-VR Training**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light- | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| headedness or spinning feeling)? | | | | | | | |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced fatigue (e.g., feeling of tiredness or sleepiness)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced discomfort (e.g., eyestrain, blurred vision, or headache)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please write below any additional comments relevant to the question above:

**Questionnaire -Training**

Please read the below statement carefully and rate how much you agree or disagree with it.

It was easy to follow the instruction to takeover from the car.

| 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please read the below statements carefully and rate how much you agree or disagree with them.

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I trust the automation in this situation. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel calm. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel nervous. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I am tense. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate your sleepiness during the scenario you've just finished.

| 1 - Extremely alert | 2 - Very alert | 3 - Alert | 4 - Rather alert | 5 - Neither alert nor sleepy | 6 - Some signs of sleepiness | 7 - Sleepy, but no effort to keep awake | 8 - Sleepy, some effort to keep awake | 9 - Very sleepy, great effort to keep awake, fighting sleep |
|---|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate, by marking the horizontal axis below how much effort it took you to complete the tasks in the scenario you've just finished.

0- Absolutely no effort
37 - 57- Some effort to rather much effort
150- Extreme effort

0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150

Please indicate

**Wait -1**

Now, you will receive further instructions from the researcher. **Please inform them that you have filled in the questionnaire and do not leave this page unless instructed to do so.**

Afterwards, when the researchers tells you, you can come back to this questionnaire and fill out the remaining questions.

**CSQ-VR-1**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light-headedness or spinning feeling)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced fatigue (e.g., feeling of tiredness or sleepiness)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced discomfort (e.g., eyestrain, blurred vision, or headache)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please write below any additional comments relevant to the question above:

**Questionnaire - 1**

Please read the below statement carefully and rate how much you agree or disagree with it.

It was easy to follow the instruction to takeover from the car.

| 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please read the below statements carefully and rate how much you agree or disagree with them.

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I trust the automation in this situation. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel calm. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel nervous. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I am tense. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate your sleepiness during the scenario you've just finished.

| 1 - Extremely alert | 2 - Very alert | 3 - Alert | 4 - Rather alert | 5 - Neither alert nor sleepy | 6 - Some signs of sleepiness | 7 - Sleepy, but no effort to keep awake | 8 - Sleepy, some effort to keep awake | 9 - Very sleepy, great effort to keep awake, fighting sleep |
|---|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate, by marking the horizontal axis below how much effort it took you to complete the tasks in the scenario you've just finished.

0- Absolutely no effort
37 - 57- Some effort to rather much effort
150- Extreme effort

| | 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 | 130 | 140 | 150 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Please indicate | | | | | | | | | | | | | | | | |

**Wait -2**

Now, you will receive further instructions from the researcher. **Please inform them that you have filled in the questionnaire and do not leave this page unless instructed to do so.**

Afterwards, when the researchers tells you, you can come back to this questionnaire and fill out the remaining questions.

**CSQ-VR-2**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light-headedness or spinning feeling)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced fatigue (e.g., feeling of tiredness or sleepiness)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced discomfort (e.g., eyestrain, blurred vision, or headache)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please write below any additional comments relevant to the question above:

**Questionnaire -2**

Please read the below statement carefully and rate how much you agree or disagree with it.

It was easy to follow the instruction to takeover from the car.

| 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please read the below statements carefully and rate how much you agree or disagree with them.

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I trust the automation in this situation. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I feel calm. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel nervous. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I am tense. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate your sleepiness during the scenario you've just finished.

| 1 - Extremely alert | 2 - Very alert | 3 - Alert | 4 - Rather alert | 5 - Neither alert nor sleepy | 6 - Some signs of sleepiness | 7 - Sleepy, but no effort to keep awake | 8 - Sleepy, some effort to keep awake | 9 - Very sleepy, great effort to keep awake, fighting sleep |
|---|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate, by marking the horizontal axis below how much effort it took you to complete the tasks in the scenario you've just finished.

0- Absolutely no effort
37 - 57- Some effort to rather much effort
150- Extreme effort

0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150

Please indicate

**Wait -3**

Now, you will receive further instructions from the researcher. **Please inform them that you have filled in the questionnaire and do not leave this page unless instructed to do so.**
Afterwards, when the researchers tells you, you can come back to this questionnaire and fill out the remaining questions.

**CSQ-VR-3**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light-headedness or spinning feeling)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced fatigue (e.g., feeling of tiredness or sleepiness)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced discomfort (e.g., eyestrain, blurred vision, or headache)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please write below any additional comments relevant to the question above:

**Questionnaire -3**

Please read the below statement carefully and rate how much you agree or disagree with it.

It was easy to follow the instruction to takeover from the car.

| 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please read the below statements carefully and rate how much you agree or disagree with them.

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I trust the automation in this situation. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel calm. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel nervous. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I am tense. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate your sleepiness during the scenario you've just finished.

| 1 - Extremely alert | 2 - Very alert | 3 - Alert | 4 - Rather alert | 5 - Neither alert nor sleepy | 6 - Some signs of sleepiness | 7 - Sleepy, but no effort to keep awake | 8 - Sleepy, some effort to keep awake | 9 - Very sleepy, great effort to keep awake, fighting sleep |
|---|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate, by marking the horizontal axis below how much effort it took you to complete the tasks in the scenario you've just finished.

0- Absolutely no effort

37 - 57- Some effort to rather much effort
150- Extreme effort

0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150

Please indicate

**Wait -4**

Now, you will receive further instructions from the researcher. **Please inform them that you have filled in the questionnaire and do not leave this page unless instructed to do so.**
Afterwards, when the researchers tells you, you can come back to this questionnaire and fill out the remaining questions.

**CSQ-VR-4**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light-headedness or spinning feeling)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience a visually induced fatigue (e.g., feeling of tiredness or sleepiness)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced discomfort (e.g., eyestrain, blurred vision, or headache)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please write below any additional comments relevant to the question above:

**Questionnaire -4**

Please read the below statement carefully and rate how much you agree or disagree with it.

It was easy to follow the instruction to takeover from the car.

| 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please read the below statements carefully and rate how much you agree or disagree with them.

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I trust the automation in this situation. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel calm. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel nervous. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I am tense. | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ |

Please indicate your sleepiness during the scenario you've just finished.

| 1 - Extremely alert | 2 - Very alert | 3 - Alert | 4 - Rather alert | 5 - Neither alert nor sleepy | 6 - Some signs of sleepiness | 7 - Sleepy, but no effort to keep awake | 8 - Sleepy, some effort to keep awake | 9 - Very sleepy, great effort to keep awake, fighting sleep |
|---|---|---|---|---|---|---|---|---|
| ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ | ◯ |

Please indicate, by marking the horizontal axis below how much effort it took you to complete the tasks in the scenario you've just finished.

0- Absolutely no effort
37 - 57- Some effort to rather much effort
150- Extreme effort

| | 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 | 130 | 140 | 150 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Please indicate | | | | | | | | | | | | | | | | |

**Wait -5**

Now, you will receive further instructions from the researcher. **Please inform them that you have filled in the questionnaire and do not leave this page unless instructed to do so.**

Afterwards, when the researchers tells you, you can come back to this questionnaire and fill out the remaining questions.

**CSQ-VR-5**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light-headedness or spinning feeling)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced fatigue (e.g., feeling of tiredness or sleepiness)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced discomfort (e.g., eyestrain, blurred vision, or headache)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please write below any additional comments relevant to the question above:

**Questionnaire -5**

Please read the below statement carefully and rate how much you agree or disagree with it.

It was easy to follow the instruction to takeover from the car.

| 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please read the below statements carefully and rate how much you agree or disagree with them.

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I trust the automation in this situation. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel calm. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel nervous. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I am tense. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate your sleepiness during the scenario you've just finished.

| 1 - Extremely alert | 2 - Very alert | 3 - Alert | 4 - Rather alert | 5 - Neither alert nor sleepy | 6 - Some signs of sleepiness | 7 - Sleepy, but no effort to keep awake | 8 - Sleepy, some effort to keep awake | 9 - Very sleepy, great effort to keep awake, fighting sleep |
|---|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate, by marking the horizontal axis below how much effort it took you to complete the tasks in the scenario you've just finished.

0- Absolutely no effort
37 - 57- Some effort to rather much effort
150- Extreme effort

0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150

Please indicate

**Wait -6**

Now, you will receive further instructions from the researcher. **Please inform them that you have filled in the questionnaire and do not leave this page unless instructed to do so.**

Afterwards, when the researchers tells you, you can come back to this questionnaire and fill out the remaining questions.

**CSQ-VR-6**

Please, from 1 to 7, choose the response that better corresponds to the presence and intensity of the symptom.

CyberSickness in Virtual Reality Questionnaire (CSQ-VR)

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| Do you experience nausea (e.g., stomach pain, acid reflux, or tension to vomit)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience dizziness (e.g., light-headedness or spinning feeling)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience disorientation (e.g., spatial confusion or vertigo)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience postural instability (i.e., imbalance)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced fatigue (e.g., feeling of tiredness or sleepiness)? | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Do you experience a visually induced | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

| | 1 - Absent Feeling | 2 - Very Mild Feeling | 3 - Mild Feeling | 4 - Moderate Feeling | 5 - Intense Feeling | 6 - Very Intense Feeling | 7 - Extreme Feeling |
|---|---|---|---|---|---|---|---|
| discomfort (e.g., eyestrain, blurred vision, or headache)? | | | | | | | |

Please write below any additional comments relevant to the question above:

**Questionnaire -6**

Please read the below statement carefully and rate how much you agree or disagree with it.

It was easy to follow the instruction to takeover from the car.

| 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please read the below statements carefully and rate how much you agree or disagree with them.

| | 1- Strongly disagree | 2- Disagree | 3- Somewhat disagree | 4- Neither agree nor disagree | 5- Somewhat agree | 6- Agree | 7- Strongly agree |
|---|---|---|---|---|---|---|---|
| I trust the automation in this situation. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel calm. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I feel nervous. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| I am tense. | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate your sleepiness during the scenario you've just finished.

| 1 - Extremely | 2 - Very alert | 3 - Alert | 4 - Rather alert | 5 - Neither alert nor | 6 - Some signs of | 7 - Sleepy, but no | 8 - Sleepy, some | 9 - Very sleepy, |
|---|---|---|---|---|---|---|---|---|
| | | ○ | | | | | | |

| alert | | | sleepy | sleepiness | effort to keep awake | effort to keep awake | great effort to keep awake, fighting sleep |
|---|---|---|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Please indicate, by marking the horizontal axis below how much effort it took you to complete the tasks in the scenario you've just finished.

0- Absolutely no effort
37 - 57- Some effort to rather much effort
150- Extreme effort

0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150

Please indicate

**Open Question**

Are there any specific issues that you experienced during the experiment that you would like to report?

Powered by Qualtrics

## Appendix E

**Prompts to read for the first instruction in each scenario:**

Please indicate how much effort it took you to complete the task from 0 to 150, 0 is absolutely no effort, 57 is rather much effort, and 150 is extreme effort.

Now, please indicate how sleepy you are from 1 to 9, 1 is extremely alert, 5 is neither alert nor sleepy and 9 is very sleepy.

Please rate how much you agree with the following statements.
*I trust the automation in this situation* from 1 to 7, 1 is strongly disagree and 7 is strongly agree.
*I feel calm* from 1 to 7, 1 is strongly disagree and 7 is strongly agree.
*I feel nervous* from 1 to 7, 1 is strongly disagree and 7 is strongly agree.
*I am tense* from 1 to 7, 1 is strongly disagree and 7 is strongly agree.

**Prompts to read on the following instructions in each scenario:**

You rated your previous effort it took you to complete the task as (X) from 0 absolutely no effort to 150 extreme effort. How much would you rate it now?

You rated your sleepiness as (X) from 1 extremely alert to 9 very sleepy. How much would you rate it now?

You rated your trust in the automation as (X) from 1 lowest trust to 7 is highest trust. How much would you rate it now?

You rated your calmness (X) from 1 lowest calmness to 7 highest calmness. How much would you rate it now?

You rated your nervousness as (X) from 1 lowest nervousness to 7 highest nervousness. How much would you rate it now?

You rated your tensity as (X) from 1 lowest tensity to 7 highest tensity. How much would you rate it now?

# Appendix F

```
---
editor_options:
  markdown:
    wrap: 72
output:
  html_document:
    df_print: paged
  word_document: default
---

#Section A. #Packages and Libraries
```{r}
require(tidyverse)
library(readxl)
library(broom)
library(ggplot2)
library(ggpubr)
library(broom)
library(AICcmodavg)
library(lavaan)
library(mediateP)
library(dplyr)
library(rmarkdown)
library(lavaan)
library(blavaan)
library(psych)
library(gridExtra)
library(semPlot)
library(car)
library(lme4)
library(flexmix)
library(loo)
library(rstan)
```

```{r}
options(scipen=999)
theme_set(theme_bw())
```

#Loading the data in excel format and transforming into data frames

```{r}
data_demographics <- read_excel('Data/demographics.xlsx') #This is the data for demographic
information
```
```

~~~
data_chatbot <- read_excel('Data/chatbot_verbal.xlsx') #This is the data for chatbot
group's verbal measurements
data_after_scenarios <- read_excel('Data/qualtrics.xlsx') #This is the data for both
group's qualtrics answers and success RATES
data_D <-as.data.frame(data_demographics)
data_C <- as.data.frame(data_chatbot)
data_Q <- as.data.frame(data_after_scenarios)
```

#Creating a success variable in data from qualtrics

```{r}
data_Q$Success <- (data_Q$I_1 + data_Q$I_2 + data_Q$I_3) / 3
```

#Take out Training

```{r}
data_C <- subset(data_C, Path != 'Training')
```

```{r}
data_Q <- subset(data_Q, Path != 'Training')
```

#Trait Anxiety

```{r}
data_D <- data_D %>%
  mutate(TraitAnxiety = rowMeans(select(data_D, starts_with("TA"))))
```

#CSQ_VR before the experiment

```{r}
data_D <- data_D %>%
  mutate(CSQ_VR = rowMeans(select(data_D, starts_with("CSQ_VR"))))
```

#Demographics

```{r}
demographics <- data_D %>%
  group_by(ID) %>%
  summarize(age = first(Age),
            sex = first(Sex),
            experience = first(Experience),
            trait_anxiety = first(TraitAnxiety),
~~~

````
            csq_vr = first(CSQ_VR),
            group = first(Group))
```

```{r}
print(demographics)
```

```{r}
summary(demographics)
```

```{r}
sd(demographics$age)
```

```{r}
sd(demographics$experience)
```

```{r}
sd(demographics$trait_anxiety)
```

```{r}
sd(demographics$csq_vr)
```

#CSQ_VR (Cybersickness) changes ##Creating an average of CSQ_VR data

```{r}
data_Q <- data_Q %>%
  mutate(CSQ_VR = rowMeans(select(data_Q, starts_with("CSQ_VR"))))
```

##Visualizing changes in CSQ_VR - ? Maybe - substitute this for three
boxplots - look for a good visualization for ?

```{r}
temp_data <- data_Q
temp_data$ID <- as.factor(temp_data$ID)

ggplot(temp_data, aes(x = Path, y = CSQ_VR, group = ID, color = ID)) +
  geom_point() +
  geom_line(size = 1) +
  geom_hline(aes(yintercept = CSQ_VR), color = "gray", alpha = 0.5) +
  labs(x = "Path", y = "Cybersickness", color = "ID") +
  scale_x_discrete(limits = c("Easy", "Moderate", "Hard"))
````

````
```

##Regression analysis on CSQ_VR and Group (Chatbot and form)

```{r}
lm(CSQ_VR ~ Group, data = data_Q)%>%
  summary()
```

##Regression analysis on CSQ_VR and Level of difficulty

```{r}
lm(CSQ_VR ~ Path, data = data_Q)%>%
  summary()
```

#State Anxiety ##First we need to reverse the first item (Calmness)

```{r}
data_C$SA_1 <- 8 - data_C$SA_1
StateAnxiety <- rowSums(cbind(data_C$SA_1, data_C$SA_2, data_C$SA_3))
data_Q$SA_1 <- 8 - data_Q$SA_1
StateAnxiety <- rowSums(cbind(data_Q$SA_1, data_Q$SA_2, data_Q$SA_3))
```

##Averaging the data from State Anxiety items into one variable

```{r}
data_C <- data_C %>%
  mutate(StateAnxiety = rowMeans(select(data_C, starts_with("SA"))))
```

```{r}
data_Q <- data_Q %>%
  mutate(StateAnxiety = rowMeans(select(data_Q, starts_with("SA"))))
```

###Regression analysis on State Anxiety and Trait Anxiety ##Qualtrics
(all groups)

```{r}
data_QD <- left_join(data_Q, data_D, by = "ID")
```

```{r}
lm(StateAnxiety~TraitAnxiety, data=data_QD)%>%
  summary()
```
````
````

````
##Chatbot verbal

```{r}
data_CD <- left_join(data_C, data_D, by = "ID")
```

```{r}
lm(StateAnxiety~TraitAnxiety, data=data_CD)%>%
  summary()
```

#Section B.

#Checking for manipulation effect (Perceived easiness x Level of
difficulty)

```{r}
lm(Easiness~Path, data=data_Q)%>%
  summary()
```

#Comparing the differences between chatbot group's verbal and form
measurements ##Seperating chatbot and form groups' form answers

```{r}
data_chatbotQ <- filter(data_Q, Group %in% c(1))
```

###Adding colums to data sheet and marking them as chatbot or form
(Within chatbot group)

```{r}
data_C$`Administration` <- 1
data_chatbotQ$`Administration` <- 0
```

####Merging two datasets

```{r}
data_CG <- bind_rows(data_C, data_chatbotQ)
data_CG$`Group` <- 1
```

#####Performing MANOVA

```{r}
````
````

````
model_chatbot <- manova(cbind(Effort, Sleepiness, Trust, StateAnxiety) ~ Administration,
data = data_CG)
summary(model_chatbot)

# Perform post-hoc tests (Wilks' Lambda) for pairwise comparisons
summary.aov(model_chatbot)
```

#Comparing the differences between chabot group's last verbal answer and
average of the verbal measurements ##Reading a new data set + adding a
StateAnxiety item

```{r}
data_final_verbal <- read_excel('Data/chatbot_verbal_final.xlsx') #This is the data for
chatbot group's final verbal answer
data_DF <-as.data.frame(data_final_verbal)
```

```{r}
data_DF$SA_1 <- 8 - data_DF$SA_1 #Inverting SA to be correct
StateAnxiety <- rowSums(cbind(data_DF$SA_1, data_DF$SA_2, data_DF$SA_3))
```

```{r}
data_DF <- data_DF %>%
  mutate(StateAnxiety = rowMeans(select(data_DF, starts_with("SA"))))
```

###Creating a new dataset for chatbot group's averaged verbal
measurements

```{r}
data_CA <- data_C %>%
  group_by(ID, Path) %>%
  summarize(across(c(Success, Effort, Sleepiness, Trust, StateAnxiety), mean))
```

####Combining datasets

```{r}
data_DF$`Type` <- 0
data_CA$`Type` <- 1
data_CA$`Group` <- 1
```

```{r}
data_CG_1 <- bind_rows(data_CA, data_DF)
```
````
````

````
#####Performing MANOVA

```{r}
model_chatbot_1 <- manova(cbind(Effort, Sleepiness, Trust, StateAnxiety) ~ Type, data =
data_CG_1)
summary(model_chatbot_1)

# Perform post-hoc tests (Wilks' Lambda) for pairwise comparisons
summary.aov(model_chatbot_1)
```

#Correlation Matrix between groups (Chatbot and form group) ##Combining
chatbot group's form and averaged verbal measurements

```{r}
data_full_ca <- data_CA # Using the average verbal measures
data_full_cf <- data_DF # Using the final verbal measures
data_full_cq <- data_chatbotQ
```

#Section D.

#Testing Hypotheses

##Model 1. This is without 'Group' as an IV.

```{r}
M1 <- "
#Direct effects
z_Trust ~ a*Path
Success ~ b*z_Trust
Success ~ c*Path

#TotSuccess=Success affected by Path and mediated by Trust
Totab:=a*b
TotSuccess:=Totab+c

#indirect effects of path (difficulty levels) on subvariables
z_MentalWorkload ~ d1*Path
z_Sleepiness~ d2*Path
z_StateAnxiety~ d3*Path

#indirect effects of trust on subvariables
z_MentalWorkload ~ e1*z_Trust
z_Sleepiness~ e2*z_Trust
z_StateAnxiety~ e3*z_Trust
````
````

```
#indirect effects of on subvariables on success
Success ~ f1*z_MentalWorkload
Success ~ f2*z_Sleepiness
Success ~ f3*z_StateAnxiety

#effects of path and trust on subvariables
Tot1:=d1*e1
Tot2:=d2*e2
Tot3:=d3*e3

#effects of path and trust on success madiated by the subvariables
Tot4:=Tot1+f1
Tot5:=Tot2+f2
Tot6:=Tot3+f3

TotSuccess2:= Tot4+Tot5+Tot6
Overallmodel:=TotSuccess+TotSuccess2

#Covariance
z_MentalWorkload ~~ z_Sleepiness
z_MentalWorkload ~~ z_StateAnxiety
z_Sleepiness ~~ z_StateAnxiety
"
```

###-- AVERAGE --### #Calculating z scores of variables

```{r}
data_c_bsem_ca <- data_full_ca %>%
  mutate(z_MentalWorkload = scale(Effort)) %>%
  mutate(z_StateAnxiety = scale(StateAnxiety)) %>%
  mutate(z_Trust = scale(Trust)) %>%
  mutate(z_Sleepiness = scale(Sleepiness))
```

#SEM Analysis ##M1 ###Chatbot group

```{r}
cg_sem_m1_ca <- sem(M1, data = data_c_bsem_ca)
parameterEstimates(cg_sem_m1_ca, ci=TRUE, level=0.95, boot.ci.type="perc")
summary(cg_sem_m1_ca, fit.measures = TRUE, rsquare = TRUE)
```

#Model Fitness (bic, dic, waic, looic) ##M1 ###Chatbot group

```{r}
fitMeasures(cg_sem_m1_ca)
```

````
#SEM Analysis ##M1 ###Chatbot group

```{r}
cg_BOOST_m1_ca <-sem(M1,data=data_c_bsem_ca, se = "bootstrap", bootstrap = 5000)
parameterEstimates(cg_BOOST_m1_ca, ci=TRUE, level=0.95, boot.ci.type="perc")
summary(cg_BOOST_m1_ca, fit.measures=TRUE, rsquare=TRUE)
```

###-- FINAL --### #Calculating z scores of variables

```{r}
data_c_bsem_cf <- data_full_cf %>%
  mutate(z_MentalWorkload = scale(Effort)) %>%
  mutate(z_StateAnxiety = scale(StateAnxiety)) %>%
  mutate(z_Trust = scale(Trust)) %>%
  mutate(z_Sleepiness = scale(Sleepiness))
```

#SEM Analysis ##M1 ###Chatbot group

```{r}
cg_sem_m1_cf <- sem(M1, data = data_c_bsem_cf)
parameterEstimates(cg_sem_m1_cf, ci=TRUE, level=0.95, boot.ci.type="perc")
summary(cg_sem_m1_cf, fit.measures = TRUE, rsquare = TRUE)
```

#Model Fitness (bic, dic, waic, looic) ##M1 ###Chatbot group

```{r}
fitMeasures(cg_sem_m1_cf)
```

#SEM Analysis ##M1 ###Chatbot group

```{r}
cg_BOOST_m1_cf <-sem(M1,data=data_c_bsem_cf, se = "bootstrap", bootstrap = 5000)
parameterEstimates(cg_BOOST_m1_cf, ci=TRUE, level=0.95, boot.ci.type="perc")
summary(cg_BOOST_m1_cf, fit.measures=TRUE, rsquare=TRUE)
```

###-- QUESTIONNAIRE --### #Calculating z scores of variables

```{r}
data_c_bsem_cq <- data_full_cq %>%
  mutate(z_MentalWorkload = scale(Effort)) %>%
  mutate(z_StateAnxiety = scale(StateAnxiety)) %>%
  mutate(z_Trust = scale(Trust)) %>%
````
````

````
  mutate(z_Sleepiness = scale(Sleepiness))
```

#SEM Analysis ##M1 ###Chatbot group

```{r}
cg_sem_m1_cq <- sem(M1, data = data_c_bsem_cq)
parameterEstimates(cg_sem_m1_cq, ci=TRUE, level=0.95, boot.ci.type="perc")
summary(cg_sem_m1_cq, fit.measures = TRUE, rsquare = TRUE)
```

#Model Fitness (bic, dic, waic, looic) ##M1 ###Chatbot group

```{r}
fitMeasures(cg_sem_m1_cq)
```

#SEM Analysis ##M1 ###Chatbot group

```{r}
cg_BOOST_m1_cq <-sem(M1,data=data_c_bsem_cq, se = "bootstrap", bootstrap = 5000)
parameterEstimates(cg_BOOST_m1_cq, ci=TRUE, level=0.95, boot.ci.type="perc")
summary(cg_BOOST_m1_cq, fit.measures=TRUE, rsquare=TRUE)
```
````
````