Passenger Response to Driving Style in an Autonomous Vehicle

by

Nicole Belinda Dillen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2019

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Despite rapid advancements in automated driving systems (ADS), current HMI research tends to focus more on the safety driver in lower level vehicles. That said, the future of automated driving lies in higher level systems that do not always require a safety driver to be present. However, passengers might not fully trust the capability of the ADS in the absence of a safety driver. Furthermore, while an ADS might have a specific set of parameters for its driving profile, passengers might have different driving preferences, some more defensive than others. Taking these preferences into consideration is, therefore, an important issue which can only be accomplished by understanding what makes a passenger uncomfortable or anxious.

In order to tackle this issue, we ran a human study in a real-world autonomous vehicle. Various driving profile parameters were manipulated and tested in a scenario consisting of four different events. Physiological measurements were also collected along with self-report scores, and the combined data was analyzed using Linear Mixed-Effects Models. The magnitude of a response was found to be situation dependent: the presence and proximity of a lead vehicle significantly moderated the effect of other parameters. Stopping events also generated a higher level of response than non-stopping events. Finally, a statistically significant association between physiological responses and self-reported scores showed that such responses could potentially be used to indicate comfort or anxiety in future adaptive systems.

Acknowledgements

I would like to thank my supervisors Prof. Krzysztof Czarnecki and Prof. Oliver Schneider for allowing me the opportunity to work on this exciting project and for guiding my motivations and execution along the way.

I thank Prof. Edith Law for the opportunity to take her class on Human AI Interaction, without which the idea for this work would never have been conceived. Thank you, Edith, for having faith in my initial idea, and for encouraging me to pursue it as thesis-level research.

I thank the members of the Engineering HCI group and Haptic Lab, for all the brainstorming sessions and super useful feedback on my thoughts and ideas.

A big thank you to my WISE lab colleagues, in particular, Atrisha Sarkar and Matt Angus, for listening to my (*extremely* rare) rants as well as helping spark new ideas and pointing out possible flaws in existing ones.

Thank you, also, to my parents, Lesley and Llewellyn, who have been so supportive despite being on the other side of the world. Thank you for being patient with me even when I was busy and irritable. I promise to write to you more now.

And finally, many, *many* thanks to my wonderful friends Marko Ilievski and Samin Khan for their unwavering support, be it academic or otherwise. Thank you for being there when I needed it the most.

Dedication

To my parents, Lesley and Llewellyn, who sacrificed all they could to make me the independent woman I am.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

With progress on autonomous vehicle technology being made at a global level, the future of autonomous vehicles is at our doorstep.

In 2018, Norway, China and Japan legalized the testing of autonomous vehicles on public roads, while Waymo, one of the leading companies in the field, started charging their passengers for rides in its vehicles. Nanyang Technological University (NTU) in Singapore collaborated with local and government bodies to set up a 2 hectare autonomous vehicle test facility; equipped with scyscrapers, a rain simulator, and even a floodzone, the facility is aimed at significantly accelerating the testing of autonomous vehicles. Meanwhile in the US, the California Department of Motor Vehicles lifted the requirement for the presence of a safety driver in vehicles being tested on public roads.

Despite these recent advancements in autonomous vehicle development, however, recent studies have indicated that the safety of fully autonomous vehicles is far from being completely accepted by the public. According to a 2019 AAA survey, 60 to 80% of drivers when interviewed expressed their fear of riding in a fully autonomous vehicle [23]. Another study based in the UK [30], evaluated social perceptions of autonomous vehicles through different road-user perspectives, i.e., driver, pedestrian, and passenger; from the perspective of passengers, such vehicles were perceived to be less risky than their human-operated counterparts. A survey conducted in Australia found that negative opinions stemmed from control, safety, and trust issues [45]. In particular, thoughts swayed around unpredictability and the possible incapability of the automated system to safely respond to emergency situations, as well as the loss of control of the vehicle and the absence of a driver to take over if necessary.

The problem with current research, however, is its tendency to focus on the safety driver, i.e., ensuring safe and timely handover, and not on the passenger. While driver handover is certainly an important topic to investigate, especially due to the current prevalence of level 2 and 3 automated driving systems (ADSs), the future of autonomous driving lies in higher level systems that do not necessarily require the presence of a safety driver at all times. Given that loss of control and distrust in the safety of the ADS is a major concern among passengers, efforts should be made to promote a more comfortable environment that promotes a sense of confidence in the passenger. This would ultimately assist in achieving the overall goal of allowing occupants of a fully autonomous vehicle to carry out their personal activities without worrying about their safety being compromised.

In order to promote this sense of comfort, however, we first need to understand what it is that makes a passenger uncomfortable or anxious. This now opens up a new area of exploration: how do different decisions taken by the ADS affect the overall well-being of the passenger? More specifically, how does the driving style of the vehicle influence the passenger's comfort or anxiety levels? This question has been previously explored through passenger-studies, but these have either been conducted in a simulated driving environment [6], or without the use of explicit physiological feedback [25].

In this thesis, we further explore this research question by executing a study in a physical autonomous vehicle. By manipulating different variables in the vehicle's driving profile, we expose passengers to different driving styles and subsequently analyze their physiological responses to the manipulations. In order to ensure our results are holistic, we also obtain, via questionnaires, their perceived levels of anxiety and comfort corresponding to the different investigated driving styles.

In the next few chapters, we discuss literature related to the driver handover task and then address the future shift in focus from drivers to passengers. We then go on to address how physiological response can be used as an indirect measure of comfort or anxiety, and describe our experimental setup and procedure. Finally, we provide a statistical analysis for the collected data, a major conclusion of which was that the presence and proximity of a lead vehicle not only raised the magnitude of all physiological responses but also moderated the existing effect of longitudinal acceleration and jerk. We also analyse the relationship between physiological responses and the self-reported comfort and anxiety scores and show how the responses are indirect indicators of passenger comfort and anxiety.

1.1 Contributions

The contributions of this work lie in the:

- collection of a naturalistic data set consisting of the physiological responses from passengers combined with corresponding vehicle state signals from a physical autonomous vehicle
- correlation of driving style parameters with physiological measurements at a coarse and detailed level of resolution
- correlation of driving style and physiological measurements with self-reported comfort and anxiety scores

Chapter 2

Literature Review

As discussed in the previous chapter, most research in human-vehicle interaction in the context of automated driving tends to focus on the human driver or –more precisely– the "fallback ready user" as defined by the SAE (in this work we continue to use the term "human driver" for consistency with existing literature).

This chapter aims to capture the state-of-the-art research directions in the field, and its necessary progression towards the inclusion of passengers and their interaction with autonomous vehicles. In addition, it addresses the issues of physiological sensing to capture comfort and anxiety as well as provides an overview of simulator studies and the need for their in-car counterparts.

2.1 Levels of Automated Driving

The SAE defines 6 different levels for an automated driving system (ADS) [15]:

- Levels 0, 1, 2 in which the driver is either completely (level 0) or partially (levels 1 and 2) responsible for the driving task
- Level 3 in which the ADS is responsible for the driving task within its operational domain but may require the occasional intervention of a human driver
- Level 4 in which the ADS is completely responsible for the driving task but only within its operational domain

- Level 5 in which the ADS is completely responsible for the driving task in any situation, i.e., without being limited to a specific operational domain.

We see that the human driver is strictly required for the first four levels while, level 4 and 5 vehicles pose no such requirement. A major task, then, for an ADS that is level 3 or below is to safely and effectively hand over control to the human driver in situations in which it is unable to autonomously make driving decisions. Such situations may occur in the event of unruly weather such as snowstorms, rain, or fog, as well as in the case of system failures.

For an ADS that is level 4 or 5, the task shifts from engaging drivers, to focusing on the passenger. In an environment without a human driver, the passenger should be able to comfortably carry out their in-car activities without any fear for the reliability and safety of the ADS.

2.2 The status quo - focusing on driver handover

For level 2 and 3 vehicles, the ADS is completely in control for extended periods of time. Consequently, the human " may become partially or completely disengaged from the driving task [16] and may instead direct their attention to tasks that were previously considered secondary in nature. The most common of these tasks involve interaction with cellphones and other handheld electronic devices [18]. In the event of an emergency takeover request, the human driver would need to be alerted before the vehicle control is completely handed over to them so as to prevent the remnants of prior work from affecting their ability to safely control the vehicle [65].

Proposed systems in current literature usually rely on multi-modal cues to communicate automotive intent to the driver. Van der Heiden et al. [73], investigated the user of audio pre-alerts to prime distracted drivers for handover requests, whereas Walch et al. [76] attempted to avoid a complete handover of control claiming that frequent handover requests could result in driver annoyance. To circumvent this, they proposed a cooperative driving solution in which an audio-visual interface was used for monitoring or approving driving decisions instead of fully handing over control to the driver.

Promoting the situational awareness of a human driver by means of spatial and navigational cues is an important pre-requisite for safe handovers. Haptic feedback in the form of vibrotactile eyeglasses and car seats [44], [68] as well as pneumatic shoulder pads and floorboards [43] have been used to provide a physical stimulus to the driver to convey directional as well as spatial information.

Such research, however, applies mainly to Level 3 vehicles that have a strict requirement for a human driver. In Level 4 and 5 vehicles, on the other hand, there is no such requirement as the vehicle itself would have a robust emergency policy in place. The question then arises: how can the shift in system design be made such that the focus now lies on passengers as opposed to drivers?

2.3 The shift in trends - from driver takeover to passenger comfort

Transitioning from semi-autonomous (Level 2-4) to fully autonomous (Level 5) vehicles will likely have a profound impact on passenger perceptions. For one, the absence of a steering wheel and pedals poses a significant change in the vehicle interior. By extension, the presence of a driver to take over control "just in case" is completely removed. With strict requirements on human drivers for Level 3 (and lower) vehicles, passengers are likely to have more trust and faith in the functioning of the vehicle knowing that a human driver would be present in the event that things go awry. However, when the human driver is taken out of the picture, passengers might be confronted with anxiety regarding their safety in a vehicle that apparently has no option for an emergency takeover.

The driving style of the vehicle itself may also be a significant cause for passenger anxiety. The driving style refers to the overall driving behavior in terms of speed, acceleration patterns, distance to leading vehicles, and traffic law obedience and is usually categorized in terms of aggressiveness or defensiveness [6]. The Multidimensional Driving Style Inventory (MDSI) [67], defines it in terms of four broad domains: (a) reckless and careless, (b) anxious, (c) angry and hostile and (d) patient and careful. The study conducted in [6], focuses on the degree of defensiveness, which is an aggregate of driving features that included mean distance to lead car, mean time headway, time headway during lane change, distance headway during lane change, distance headway merge back, braking distance from the intersection, time to stop, maximum turn speed, speed at intersection, and average speed for 20 metres before the intersection.

Work that deals with learning driving styles typically involves modelling the individual style of the human driver. Abeel and Ng [1] proposed the use of inverse reinforcement learning (IRL) to optimize an unknown reward function using a linear combination of features, which can, in turn, be used to solve the task of learning the driving style. Kuderer et al. [36] also proposed a learning by demonstration approach in which they model the driving style using a cost function that can then be optimized using feature-based IRL.

Other works that use IRL to model driving styles include Babes et al. [5] and Ziebart [86] who use maximum likelihood estimation variants of the technique to model simpler highway driving.

While IRL is a common approach to modelling driving style, there also exists supervised learning solutions to the issue. Dong et al. [20] trained several deep learning models (consisting of Convolutional Neural Network and Recurrent Neural Network variants) to characterize driving styles from GPS sensor data. Xu et al. [82] use a partly connected multi-layered perceptron (PCMLP) to imitate human driving for the purpose of vehicle testing in industry.

Most work suggests that passengers would prefer a driving style based on their own: aggressive if the passenger drives aggressively, and more defensive otherwise. Basu et al. [6], however, showed that most passengers actually prefer to be driven according to their perceived driving style which is usually more defensive compared to their own. This discovery undermines work that relies on identifying the passenger's personal driving style through techniques such as learning by demonstration [36]. The new task is, therefore, to accommodate the passenger's discomfort with the current driving style and alter it accordingly.

Situations might also arise in which the passenger is perfectly comfortable with the driving style itself but may need to stop or change the route. For example, the option to take a slower but more scenic route might be available – should the car continue on the quicker but less scenic path? Perhaps the car might detect a stationary object up ahead – could this be an accident and the passenger the first to arrive on the scene? What about passengers with motion sickness [63] or disabilities [19]? For them, it would be beneficial to look into different ways of promoting situational awareness and communicating future navigation decisions.

Systems that can circumvent these new challenges propose to keep the user in the loop, and though they are currently targeted at drivers, they are equally applicable for use by passengers. For instance, Stewart [69] is a haptic device that allows the user to engage in a "haptic discussion" with it. It can convey upcoming driving decisions through subtle movements, and allows the user to override (if legal) the decision by changing the direction of movement of the device. "Scribble" [54] is another interface that allows the user to change driving decisions by "drawing" their way through traffic. Toyota also demonstrated a prototype – dubbed "Yui" – of a human-like artificial intelligence based user interface that is able to converse with drivers and passengers and learn their preferred routes [52].

The drawback with systems like Stewart and Scribble is that they require the passenger to explicitly intervene and change the driving decision taken. Thus, besides being anxious about the current driving style, they would also need to pause their current in-car activities in order to manually change the driving decision through an interface. Our solution to this is to propose an affect-aware hardware system that does not rely on identifying a users preferred driving style beforehand. The system would instead be responsible for sensing vehicle state data as well as physiological data from the passengers in order to determine their level of anxiety. Using the predicted anxiety level, the system would then be able to automatically adjust for a more appropriate driving behavior that would reduce the passengers anxiety. In addition, future decisions would simultaneously be communicated to the passenger via audio-visual and haptic feedback. Through this sense-feedback-sense loop, the system would enable passengers to engage in their personal in-car activities uninterrupted, while responding to their anxiety levels and promoting a sense of confidence in the vehicle's autonomy.

This thesis is a natural consequence of the first phase of this solution, i.e., to understand the physiological response of a passenger to different driving styles in order to determine the onset of discomfort or anxiety. The sensing should ideally be done in a non-invasive manner and should not require any explicit input from the passenger.

Our work, therefore, falls under the realm of implicit HCI which we discuss immediately in the next section.

2.4 Implicit HCI

Implicit HCI is an emerging area of HCI research that allows a smart system to perform actions based on the environment and context and without explicit intervention from the user. Schmidt [60] defines it as "an action performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as an input". Through it, a system can garner the power of sensing technology to build its awareness of context and user affect and perform an action or provide a solution in an anticipatory manner, rather than having to rely on an explicit cue from the user. Such systems fit well into the bigger picture of ubiquitous computing, a term coined by Weiser [78], who described a future in which people and their environment would be augmented with computing technology to allow for seamless, naturalistic interactions that serve to provide services and information or even alter the surrounding environment.

Schmidt [60] proposed the development of an implicit HCI system through sensing the context around a user. An XML-based mark-up language was used to specify context and action trigger semantics for a system (based on whether the system is entering, inside or leaving a context). Proposed methods for gathering contextual data were also discussed and included a speech and position awareness wearable built into a tie, as well as a sensor board consisting of a touch sensor, a passive infrared sensor, an accelerometer, a light sensor and a temperature sensor. Schmidt and Merz [61] went on to propose an RFID wearable that allowed the user to interact with tagged RFID objects implicitly; for example, they explored the use-case of real-world bookmarks in which interacting with an everyday object triggers the execution of a computer application.

Zander et. al. [83] devised an implicitly controlled brain-computer interface that used EEG signals from a user to passively control the movement of a cursor on a screen. The users in this study did not actively control the movement of the cursor but simply watched and interpreted its initially random movements, after which the cursor responded to this interpretation (acquired via the EEG signals) by decreasing the randomness in its movements.

In our case, the implicit system consists of vehicle state variables that make up the context of the user's (passeneger's) environment. Vehicle state variables such as acceleration, jerk, and stopping distance can be measured through sensors installed on-board. Wearable sensors such as skin response and heart rate sensors, as well as eyetracking glasses make up the physiological signal streams that, when combined, represent the affective state of the user.

In order to properly construct our system and study design, we must first define the term anxiety and determine how it can be appropriately measured.

2.4.1 Measuring Anxiety and Discomfort

In this work, we aim to predict the presence of anxiety, or more generally, discomfort, in passengers in autonomous vehicles. More specifically, we study state anxiety which has been defined by Spielberger [64] as a complex emotional response to perceived threat characterized by "feelings of tension and heightened autonomic nervous system activity".

Eyetracking

Anxiety has been linked to eye movement patterns; in particular, loss in focus on the areas of interest seems to be a trend among individuals experiencing anxiety. Eyetracking research is primarily based on the concept of gaze points and fixations. Gaze points are simply the raw samples captured by the eye tracking device at the sampling frequency.

Fixations, on the other hand, represent clusters of gaze points that occur in quick succession (typically within 100-300 milliseconds) and within a small range of space.

It has been reported that, in conditions of high state anxiety, attention control appears to be more stimulus-driven as opposed to goal-driven [79]. In competitive environments, this results in sub-optimal behavior in which more time is spent searching for goal-irrelevant cues in the environment instead task relevant cues. Evidence of stimulus-driven control lies in the associated patterns of gaze behavior which is characterized by a decrease in the duration of goal-oriented fixations as well as an increase in the number of fixations.

In the context of driving, Fridman et al. [37], [26] used vision-based gaze estimation for driver monitoring, particularly to predict driver state and behavior, such as fatigue and failure to signal. Wilson et al. studied gaze behavior in the context of anxiety, attention, and penalty kick performance [80] using fixation metrics that included time to first fixation, total fixation duration, total number of fixations at the target areas (the goal area and goalkeeper). Participants were found to make significantly more fixations to both target areas under high-threat situations and spent a longer duration focusing on the goal keeper as opposed to the goal target area resulting in sub-optimal shooting strategies.

Causer et al. [12] explored the relationship between anxiety and performance through gaze behavior in the context of competitive shotgun shooting. High anxiety conditions resulted in shorter final fixations on goal areas reflecting disruption to attentional control processes during such conditions. This was further corroborated by Allsop et al. [3], who found that high anxiety increased scanning entropy - defined as the randomness or predictability of scan behavior - in the presence of cognitive load. They also found that, in contrast to the low anxiety condition, the percentage dwell time on the external world in the high anxiety condition was significantly higher, accompanied by a significant decrease in the percentage dwell time on the area of interest.

Heart Rate and Heart Rate Variability

State anxiety is commonly measured through autonomous system variables such as heart rate (HR) and heart rate variation (HRV).

The effects of parasympathetic activity in a state of anxiety were studied in [42]. A power spectral analysis was performed on the ECG signal of a participant in order to obtain the frequency decomposition of the HRV. Results showed that the measure for low frequency (LF) components was found to be higher in a state of anxiety as compared to baseline conditions.

Morales et al. [40] studied the sensitivity of HR and HRV to anxiety caused by stressful situations before judo competitions. They found that LF components were higher both in more stressful official (as opposed to unofficial) competitions as well as among less-confident and more anxious national-level players (as compared to more confident international players). Similarly, in the more experienced international players the high frequency (HF) band was higher and the HR was lower.

Rani et al. [51] exploited these effects of parasympathetic activity (in a state of anxiety) on HRV by performing a power spectral analysis on the inter-beat interval obtained from a photoplethysmogram (PPG) signal. Spectral features along with skin conductance response and facial cues were then used as inputs to a fuzzy logic system to produce an anxiety index for a robot interacting with a human. In [38], among other physiological measures, the mean and standard deviation of the inter-beat interval and peak PPG amplitude was used along with the low frequency power obtained from the ECG signal to train a regression tree classifier.

PPG features (along with EEG features) were used to train a kNN and SVM classifier to predict anxiety in [85]; this study also found that the mean pulse rate could be a potential marker for anxiety. However, contrary to the studies mentioned above, the authors did not find a significant cause-effect relationship between HRV features and anxiety. In particular, while their experimental results for low and high frequency powers generally agreed with the trends seen in literature, their actual correlation with anxiety levels was low. They speculate that this might be due to both individual differences in physiological baselines as well as the choice of time span adopted for the frequency decomposition of the HRV.

Galvanic Skin Response

In addition to HR and HRV, Galvanic Skin Response (GSR) is another common autonomous nervous variable that has been found to be a significant indicator of anxiety. GSR variables commonly include the Skin Conductance Level (SCL) and the Skin Conductance Response (SCR). The former, SCL, is a measure of tonic activity and changes slowly over time, usually over a period of seconds or even minutes. SCL, on its own, is typically not too informative as it can significantly differ across individuals and is, therefore, usually evaluated in terms of a baseline level.

SCR is a measure of phasic activity and is characterized as quick bursts of elevated conductance levels. These quick bursts are visualized as GSR "peaks" as are highly sensitive to emotionally arousing events. Event related SCRs (ER-SCRs) typically occur 1-5 seconds from the onset of emotional stimuli. Some GSR peaks, on the other hand, occur irrespective of any emotional stimuli. These non-specific SCRs (NS-SCRs) may manifest spontaneously at a rate of 1-3 per minute. A single event stimuli may generate multiple successive peaks.

Alpers et al. [4] conducted a study on driving phobic patients and found significant effect sizes in SCL during exposure to the anxiety-inducing driving task as well as in pre-exposure and post-exposure conditions. In fact, elevated SCL was also observed among the non-phobic control participants, although these values were lower compared to those seen in phobic patients. Similar effects for SCL were seen by Blechert et al. [9], who studied the effects of peripheral physiological information on anxiety and found the largest effect sizes demonstrated by electrodermal measures. Besides SCL, these measures also included the NS-SCR rate and SCR amplitude.

Increased skin conductance levels (measuring tonic activity) were also shown to be an indicator of anxiety in studies on attention manipulation [42] and respiratory resistance in asthma patients [53], as well as in a study on expectancy bias in trait anxiety [13].

Skin conductance response (measuring phasic activity) was another major indicator of anxiety in the context of dynamic difficulty adjustment in computer games [38], a study which based its findings on the results of previous research on affect recognition in human-robot interaction [51]. In both studies, a valid response was one whose slope of the rise to the peak was greater than 0.05 micro-Siemens/minute with an amplitude greater than 0.05 micro-Siemens and a rise time greater than 0.25 seconds. The number of valid responses were calculated in intervals to determine the rate of response along with the mean phasic amplitude of all responses. Using these metrics, it was found that both the mean amplitude of the SCR as well as the rate of responses increased with an increase in anxiety.

In another similar study by the aforementioned authors [50], electrodermal activity was studied in the context of human-machine interaction and included variables such as the mean tonic and phasic activity, rate of phasic activity, as well as the slope of tonic activity. As expected, the mean tonic and phasic activity levels were always positively correlated with anxiety.

2.5 Simulated versus real-world autonomous vehicle environments

In order to circumvent the difficulties and safety concerns of using a real-world vehicle, most autonomous vehicle user studies involving drivers and passengers have been conducted using driving simulators. This is largely due to both the limited risk of using a driving simulator as well as the ease of data collection and customization of specific vehicle and environmental features and scenarios. For instance, a major advantage of using a simulator is the ability to test and program crashes into the study environment, a situation that, for obvious reasons, can never be tested with users in a real-world setting. The in situ nature of a simulator study also allows for greater overall control in the experiment by allowing the experimenter to tweak only specific scenario settings without affecting the rest.

Driving simulators range from the more simple and low-cost set-ups that can be run on a PC such as City Car Driving [17] and Grand Theft Auto V [57] environments to highly advanced immersion systems that include Ford's Virttex simulator [28] and the Toyota Driving Simulator [72]. For most researchers, however, the more realistic and advanced systems are out of reach, leading to a compromise between naturalistic and simulated driving experiences.

Simulator results have been found to be transferable to the real world with respect to driving experience. Visual attention and task performance were found to be relatively similar between an on-road and medium-fidelity simulator studies [77]. Straight road driving was also found to be comparable between a simulated environment and real-world driving [8]; however, lateral control was affected by the absence of kinesthetic feedback resulting in a diminished perception of lateral translation. In general, it has been found that while simulator driving is able to approximate in-car driving, it is unable to replicate it, making it unsuitable for use-cases in which absolute validity is required [41].

Thus, while the benefits of a simulated environment clearly warrant their continued use in future work, the ability to capture human responses in an actual real-world vehicle is incredibly useful. In-car studies contribute not just to the external validity of an experiment but also generate usable real-world data. The additional risk involved will also capture the full attention of participants who otherwise might take a simulator study for granted.

Efforts have been made to bring low-cost simulated driving as close as possible to reality. In [58], simulated driving was coupled with motion effects in order to investigate the effect of motion on take-over requests, while in-car virtual reality was used in [27] to map physical car motion to a virtual environment.

Most autonomous vehicle studies in literature rely on the use of driving simulators to study take-over requests. The STIsim simulator was used in [73] along with a Ford dashboard, steering wheel, acceleration and break pedal to test the effectiveness of pre-alerts in priming drivers for handover. Driving simulators were also used in a study on voice alerts for distracted drivers [81], and to evaluate various auditory displays for take-over requests [32].

On-road studies have been performed mainly in the context of pedestrian interactions with autonomous vehicles either using a Wizard-of-Oz approach to automated driving [55], or the actual ADS itself [74].

With respect to passengers, however, simulator studies are still prevalent: a highly immersive simulator-based study was performed to understand the passenger's experience of working with full autonomy engaged [49], while driving style preferences were evaluated with a low-fidelity simulator consisting of a classroom projection screen coupled with a Logitech G920 steering controller [6]. An in-car study was, however, performed by Festner et al. [25], who investigated passenger comfort in response to different lane-changing behaviors.

Chapter 3

Study

We conducted a study on a closed test-track located in Waterloo, Ontario, Canada. This chapter highlights the agents involved in the study, the research questions and variables involved, the proposed study design and rationale behind it, as well as the recruitment procedure and participant demographics.

3.1 Agents involved

The study involved the use of an ego vehicle (the autonomous vehicle in which the participant was seated) and one other agent vehicle driven by a human. There were no other vehicles or pedestrians present in the study design.

The **ego vehicle** used in this study was the Autonomoose, a Lincoln MKZ hybrid used as a research platform for autonomous driving. It is fitted with lidar, inertial and vision sensors and is currently equipped with Level 3 driving autonomy. Although the study was run completely with autonomy engaged, a safety driver was present at all times to take over control if necessary.

The **agent vehicle** was a black Lexus 450 Rx - this vehicle was always driven manually and had no autonomous capabilities.

3.2 Research Questions

We wished to answer the following research questions:

- *RQ1:* How does driving style correlate with passenger physiological response?
- *RQ2:* How does passenger physiological response correlate with self-reported comfort and anxiety?
- *RQ3:* How does driving style correlate with passenger self-reported comfort and anxiety?

While there are several factors that contribute to the overall driving style of the vehicle, the factors we chose to control were the maximum acceleration of the ego vehicle and the minimum stopping distance it could maintain with respect to the agent.

3.3 Variables

There were a number of dependent and independent variables that made up the analysis.

3.3.1 Independent variables

At a high level, the main independent variables for the study consisted of the lateral and longitudinal components of acceleration of the ego vehicle and (the lateral and longitudinal components) of its distance to the dynamic object, i.e., the agent vehicle. Secondary independent variables were the derivatives of the acceleration components: lateral and longitudinal jerk. Of course, these variables could not be directly manipulated but were a direct result of the manipulations of the acceleration variables.

We manipulated the thresholds for both maximum acceleration and minimum distance. This manipulation was applied simultaneously to both the longitudinal and lateral components of acceleration and distance, so that at any given time, and for any variable, both components had the thresholds set to either the upper or the lower limit.

For the actual analysis, the values for these variables varied depending on the level of analysis carried out: for some analysis models the mean values for each variable were considered over a specified interval, while other analysis models considered the maximum values of each variable for the interval.

Table 3.1: Thresholds for the driving parameters varied, i.e., maximum acceleration and minimum distance. Long. and Lat. indicate the longitudinal and lateral components of the parameter in question.

Parameter	**Defensive**		**Aggressive**	
	Long.	*Lat.*	*Long.*	*Lat.*
Acceleration	8 m/s^2	4.5 m/s^2	10 m/s^2	6 m/s^2
Distance	10 m	4.5 m	7.5 m	2 m

3.3.2 Dependent variables

At a high level, dependent variables included the self-reported anxiety and comfort scores, and the physiological sensor streams:

- Galvanic skin response (GSR) variables, i.e., skin conductance level (SCL) and number of GSR peaks
- Heart Rate (HR)
- HRV variables, i.e., low and high frequency components (LF and HF) and their ratio
- Eyetracking variables, i.e., eye movement entropy, and left and right pupil dilation

As with the independent variables, both mean and maximum values for these variables were considered depending on the model used for the analysis.

3.4 Design and Scenarios

The study design proposed was aimed at isolating the different driving style parameters as much as realistically possible (even though in reality the effects of these parameters nearly always occur in conjunction).

3.4.1 Rationale for the proposed design

Initially, we chose a 2x2 within-subjects design for our study with within-subject factors of maximum acceleration and minimum distance, each with two distinct levels of threshold for both their lateral and longitudinal components, indicated in Table 3.1

As mentioned previously, thresholds were applied to both lateral and longitudinal components of each variable simultaneously. While the study could have been carried out solely on the resultant acceleration and Euclidean distance variables, it would be far more interesting as well as beneficial to actually study the effect of the components themselves. However, individually manipulating each component would result in a 4 factor study with 2 levels for each factor. This would have required 16 trials of the same driving route, which would be have been unfeasible to run given the sheer number of participants that would be needed.

Although the thresholds were manipulated, the individual parameters themselves continuously vary during realistic driving. Thus, a factor-based study design would not realistically be possible. As a result, we performed a regression analysis on the continuous values of the entire signal obtained for each of the parameters.

In order to gain a better understanding of the contribution of each component without sacrificing the simplicity of the study design, we decided to split each trial into four events, which we discuss next.

3.4.2 Teasing apart components with events

In order to tease apart the effect of each component, each trial consisted of four events:

1. A "passing" event, in which the ego vehicle was made to approach a parked agent vehicle from the left, where the agent vehicle is oriented perpendicular to the ego at the intersection. This event was meant to test the effect of lateral distance.

2. A "vehicle-stop" event, in which the ego vehicle was made to approach a parked agent vehicle from behind, and stop behind it. This event was meant to test the effect of longitudinal distance.

3. A "turn" event, where the ego vehicle was made to take a sharp turn. This event was meant to test the effect of lateral acceleration.

4. An "intersection-stop" event, in which the ego vehicle was made to come to a complete stop at an intersection (no other agent vehicle was present). This event was meant to test the effect of longitudinal acceleration. In some cases, however, the participants were able to see the presence of the agent vehicle ahead (when this event preceded the car-stop) even though there was no close stop behind the vehicle for this event (Trial 1 in Figure 3.4.2).

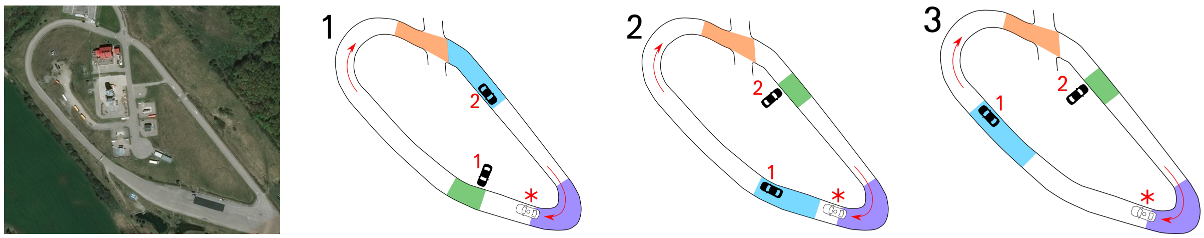


Figure 3.1: The layout of the test track. The order of events for each trial, is indicated by the number beside the vehicle and the star represents the start and end location of the ego vehicles. Colored segments of the track represent the locations for the different events: green for passing, orange for intersection-stop, blue for car-stop, and purple for turning events. Trial 4 was a repetition of trial 1.

Varying the order of events

In order to minimize boredom and keep the participants guessing, each trial had a different ordering for the events. This order was the same for all corresponding trials. It should be noted that while the trials themselves followed the same order for all participants, the order in which the *thresholds* were varied for each trial was randomized for each participant.

Since the track consisted of only one sharp turn and one intersection, the locations for the turn and intersection-stop events were fixed and could not be varied. Two locations were possible, each on opposite sides of the track, for the passing event, and three locations were possible for the vehicle-stop event (Figure 3.4.2).

Since it was necessary to have the events as far apart in timing as possible, only three unique variations in event ordering were possible. Thus, the four trials followed the order sequence (1-2-3-1), where the ordering for the fourth trial was the same as that of the first.

3.4.3 Introducing a task - watch a video

Eye movement entropy is a measure of the attention to a particular task - anxiety or apprehension can decrease attention span and lead to an increase in the randomness of eye movement patterns, i.e., the entropy. Thus, in order to measure the eye movement entropy, we introduced a task to capture and measure the attention (and indirectly, the anxiety or discomfort) of participants.

The task involved watching a video on a phone mounted to the dashboard. The video was chosen as carefully as possible to minimize any emotional response that could be elicited from watching it. It was also regarded as a more natural in-vehicle activity that does not incur much overhead.

3.4.4 Measuring individual response

Pre-study demographic questionnaire

Participants were asked to fill out a pre-study demographics questionnaire. Besides the standard questions on age and sex, the questionnaire also included questions on perceived and preferred driving style, background and experience with autonomous vehicles, and overall trust in autonomous vehicles. The questionnaire can be found in the Appendices section.

Self-reported scores

At the end of each trial, participants were asked to complete the Somatic State Anxiety sub-scale of the Competitive State Anxiety Inventory-2 questionnaire [39]. In addition to measuring the overall anxiety level of a trial, on-the-fly questions during the trial were also asked. Specifically, participants were asked two questions at the end of each event:

1. How comfortable did you feel on a scale of 1-10, 1 being "Not comfortable at all" and 10 being "Very comfortable"?
2. How anxious did you feel on a scale of 1-10, 1 being "Not anxious at all" and 10 being "Very anxious"?

The rationale behind asking these on-the-fly questions in addition to the questionnaire at the end of each trial, was to gain a better understanding of how each event influences the overall comfort or anxiety levels, and minimize the effects of the peak-end rule. Furthermore, comfort was included as an on-the-fly response measure after running a pilot study during which some participants explicitly mentioned feeling uncomfortable with the driving as opposed to anxious. A standardized questionnaire for comfort that was suitable to our study was not found and, therefore, such a questionnaire was not included at the end of each trial.

Figure 3.2: A participant sitting in the Autonomoose and set up with the eye-tracker, skin reponse, and heart rate sensors.

Physiological response

Physiological responses were measured using eye-tracking glasses, an optical ear clip, and a skin response sensor. Figure 3.2 depicts a participant set up with these sensors.

The Tobii Glasses 2 Pro were used to track gaze behavior at a sampling rate of 50Hz. Gaze points were converted into fixations using the iMotions application [47]. The Shimmer3 GSR+ device was used to collect GSR and PPG signals. GSR samples were collected by means of two electrodes fitted to the index and middle fingers of the non-dominant hand. An optical ear clip was used to obtain a PPG signal which was then processed to obtain the HR and HRV signals described earlier. Participants were instructed to keep as still as possible and to avoid speaking except when necessary, so as to not disrupt the quality of the PPG and, especially, the GSR signals. Both PPG and GSR signals were sampled at 512Hz.

3.5 Recruitment and Procedure

We wanted to recruit as diverse a group of participants as possible in order to capture differences in age, technical background, and gender.

3.5.1 Participant Recruitment

Participants were recruited from a diverse pool of volunteers consisting of individuals varying in age, familiarity with autonomous vehicles, and driving proficiency. Recruitment was done primarily through posters displayed at various departmental locations on the University of Waterloo campus. Volunteers were promised a remuneration of 15 CAD for participation as well as travel reimbursement between the test track and university campus. Interested volunteers who wore glasses were asked to either wear contact lenses during the study or be able to see decently without glasses. This was necessary in order to obtain accurate reading from the eye-tracker. Besides the requirement for no glasses, there was no other exclusion criteria.

In total, 29 participants were recruited. Of these, 1 participant's data was discarded due to technical problems that occurred during the experiment. Of the remaining participants, the first 8 (5 M, 3 F) participated in the pilot study and ranged between 20 to 46 years of age (M=22.87, S.D.=2.79). The remaining 20 participants (10 F, 10 M) participated in the main study and were aged between 19 to 64 years (M=33.5, S.D.=3.52).

3.5.2 Study Procedure

To account for ordering effects and minimize learning, participants were balanced in a randomized order.

Upon arrival at the test track, each participant was verbally briefed on the study and provided a more comprehensive information letter. After signing the consent form, participants were seated in the vehicle and fitted with the physiological sensors. They were then given the pre-study demographics questionnaire followed by an initial CSAI-2 questionnaire.

The study then commenced with the participants experiencing four trials (for each of the factor manipulations). As a source of qualitative information, participant reactions (both verbal and facial) were recorded during the course of each trial. In addition to the questionnaires, participants were interviewed at the end of the study. The interview questions were meant to understand the participant's self-described experience with the overall experimental setup and involved questions such as "Which event would you say made you the most uncomfortable or anxious?" and "Did you feel any motion sickness?". Any additional information volunteered by participants was also noted.

Chapter 4

Signal Processing

In this chapter, we discuss the Autonomoose platform used in our study, followed by the procedure in which the raw vehicle state and physiological signals were extracted and processed in order to obtain the variables used in the analyses.

4.1 The Autonomoose platform

The ADS software used in the Autonomoose consists of four major components: perception and mapping, planning, control, and supervision, each of which are further described below.

4.1.1 Perception and mapping

The **perception layer** uses Aggregate View Object Detection (AVOD) [35] to detect dynamic objects (vehicles, pedestrians, and bicycles) by consuming RGB camera images along with LIDAR data. The 3D bounding boxes for the detected dynamic objects and their orientation are then continuously passed into a Kalman Filter which estimates the speed and position tracks of the objects. Static objects (such as trees, poles, traffic signs, etc.) in the immediate surrounding area are detected using LIDAR data and are represented in a probabilistic 2D occupancy grid.

Sensor fusion is used by the **localizer** to combine LIDAR point cloud data with GPS/INS measurements and all four wheel speed measurements in order to obtain accurate position, orientation, and velocity estimation for the ego vehicle.

The **map server** produces a map of lanelets, which are essentially "atomic, interconnected drivable road segments" [7] with a polyline of GPS waypoints representing the left and right boundaries of the lane segment as well as the centreline. In addition, the map server produces the route of lanelet IDs between input start and end points.

4.1.2 Planning

Goal points are provided by the **mission planner** to the map server in order to obtain a route for the vehicle.

The route is then sent to the **behavior planner** which predicts future trajectories for dynamic objects and uses this along with the occupancy grid information to generate maneuvers for the vehicle. Such maneuvers include decelerating to a stop, maintaining speed, and yielding to other vehicles.

Maneuvers are communicated to the **local planner** in conjunction with other behavior planner output such as stop locations, lane boundaries, and dynamic objects of interest. The local planner uses this information along with occupancy grid and vehicle state information (from the localizer) to plan a smooth trajectory between the current point on the obtained path and the next point. In addition, it outputs a safe velocity profile for the car along the path.

4.1.3 Control

The **vehicle controller** executes the smooth trajectory and velocity profile generated by the local planner using a Stanley steering controller [70] for lateral control and a PID controller for longitudinal control and outputs the throttle, brake, steering, and gear positions.

4.1.4 Supervision

The **system supervisor** is responsible for managing the current state of the system, for example, whether the vehicle is in manual mode, or whether autonomy has been engaged. It also manages intervention requests and the fallback procedure followed by the system in case of an emergency.

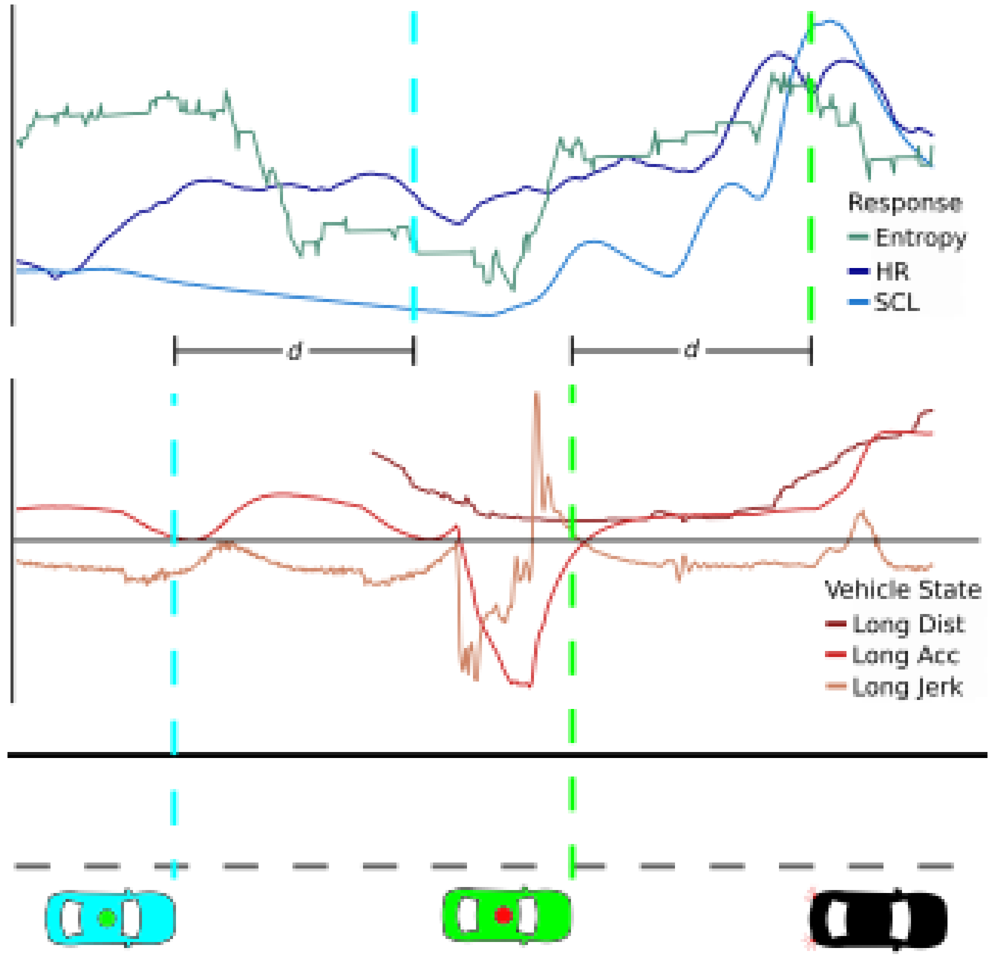


Figure 4.1: A car-stop event with physiological (above) and vehicle state (below) signals. There is an inherent and unavoidable delay, *d*, associated with the physiological signals. Due to the different scales of each signal, units are not displayed; instead, the signals were normalized in [-1,1].

4.1.5 Vehicle state signal streams

The vehicle state signal produced by the localizer contains the acceleration and jerk streams that comprise the motion-related variables used for the model, while the lateral and longitudinal distances to the dynamic object are obtained from the perception layer's tracker.

These signals are sampled at a rate of 20 Hz. Figure 4.1 shows sample longitudinal acceleration, longitudinal jerk, and longitudinal distance signals from a single participant's trial. For sample signals from all vehicle state streams, please refer to the Appendices.

4.2 Processing Physiological Signals

Figure 4.1 shows a set of sample SCL, HR, and entropy signals from a single participant. The raw signals were further processed in order to obtain the physiological variables needed for the analysis.

4.2.1 Extracting galvanic skin response variables

The raw SCL signal is obtained from the GSR Shimmer sensor itself. While the raw SCL can be (and is) used on its own, other more useful variables come from the phasic component of the GSR signal. In particular, we are interested in extracting the peaks and the maximum peak amplitude.

Before extracting the phasic component of the signal, the signal must first be pre-processed. The pre-processing step essentially involves de-trending the signal in order to make it stationary. During the experiment, due to the nature of the velcro-straps fixing the electrodes to the fingers, the amount of sweat produced increased over time causing a linear trend in the skin response signal that prevented its proper analysis. The de-trending was carried out by fitting a linear least-squares regression line to the signal and subtracting this line from the signal at corresponding points.

Extracting the phasic component

In order to extract GSR peaks and, consequently, the maximum peak amplitude, the phasic component of the raw (de-trended) signal obtained over the entire trial is first computed [31]. This is done first by filtering the GSR signal using a median filter centered on a sliding window of 9 s. For a sampling rate of 512 Hz, this would result in a window size of 4608 samples, wherein the middle sample is replaced by the median value of all samples in the window. Next, the average value of all the samples in the current 9 s window is subtracted from this median value. By sliding the window forward by one sample, and repeating this process at each step, the final phasic component of the entire signal can be obtained.

Extracting peaks

The extracted phasic component may then be used to compute peak onsets and offsets. Peak onsets were calculated by looking for all points at which the amplitude of the phasic component was greater than $0.01\mu S$. Similarly, offsets were detected by finding all points at which the phasic amplitude was less than $0\mu S$. Peaks were detected as the maximum amplitude in the de-trended signal between each pair of corresponding onsets and offsets.

4.2.2 Calculating heart rate and heart rate variability

Heart rate is quite simply the number of times the heart beats in a minute and is measured in BPM (beats per minute), while heart rate variability or HRV is defined as the variation in the interval between consecutive heartbeats. In an electrocardiogram (ECG) signal, this would be defined as the variation in the R-R intervals of a heartbeat. In a photoplethysmographic (PPG) signal (which is the signal used in this work), the variation is measured in terms of what is called the inter-beat interval or IBI.

Each cardiac cycle results in a pressure pulse propagated through the blood vessels as the heart pumps blood. This pressure pulse, though dampened, is large enough to distend the walls of the arteries and arterioles in the subcutaneous tissue resulting in changes in the volume of blood in the vessel at the point of distension. A PPG signal measures the change in this volume optically, through the use of a pulse oximeter. The skin is illuminated using a light-emitting diode and the change in blood volume is detected by measuring the amount of light transmitted (or reflected) to a photodiode.

As the PPG signal is sensitive to movement, the measurements are usually taken at relatively stationary locations on the body, most commonly the earlobe and finger.

Heart rate can be fairly straightforward to measure from the PPG signal. The raw signal - over an interval of 10 seconds - is first passed through a peak finding algorithm that searches for all local maxima by comparing the amplitudes of neighboring samples. From among these local maxima, only those with a prominence greater than 1.5 times the standard deviation are chosen as peaks of interest. The number of such peaks in the 10 s interval are then counted and multiplied by 6 in order to estimate the total number of peaks in one minute. This estimation is then recorded as the heart rate.

A series of heart rate measurements for all 10 s (overlapping) intervals over the entire trial is then obtained and is further processed in the frequency domain in order to obtain the HRV components. The HRV components, however, are measured only for the entire span of the trial, which is approximately 3 minutes in length for an individual. This is because HRV requires the measurement of very low frequency components down to 0.04 Hz - a resolution that is not quite possible with small time intervals.

The Fourier transform

The Fourier transform is a mathematical tool that is used to transform a signal in the time domain to the frequency domain. The underlying idea lies in the fact that any signal can be decomposed into a sum of sinusoidal basis functions, each with a different frequency. The Fourier transform may be applied to both periodic as well as non-periodic functions and is given by the formulae:

$$F(k) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ikx}dx \tag{1}$$

$$f(x) = \int_{-\infty}^{\infty} F(k)e^{2\pi ikx}dk \tag{2}$$

where k is the frequency being considered, $F(k)$ is the amount or "power" of the frequency k in the signal, x is the time being considered, and $f(x)$ is the value of the signal at time x. Here, (1) is called the forward Fourier transform and (2) is called the inverse Fourier transform.

The equations above can only be applied to a continuous signal and, hence, represent the continuous Fourier Transform. In real-world experiments, however, signals are usually sampled at regularly spaced intervals. We thus have only samples of the function as opposed to the actual function itself. In such a situation, the continuous Fourier Transform cannot be used and should instead be replaced with the Discrete Fourier Transform (DFT), the equations for which are:

$$F_k = \sum_{n=0}^{N-1} f_n e^{-2\pi ikn/N} \tag{3}$$

$$f_n = \frac{1}{N}\sum_{k=0}^{N-1} F_n e^{2\pi ikn/N} \tag{4}$$

where N is the number of samples, n is the sample being considered, f_n is the value of the signal at time n, k is the frequency being considered and F_k is the power of the frequency k in the signal.

Determining HRV from a raw PPG signal

HRV statistics can be computed in either the time domain or the frequency domain. Time domain statistics, however, have the disadvantage of being unreliable as well as requiring a large sample size to be computed. Frequency domain analysis, on the other hand, requires a much smaller interval of samples and is also more widely accepted.

Frequency domain analysis can be conducted by transforming the IBI series into the frequency domain through methods such as wavelet decomposition and the Fourier transform. In this work, the Fourier transform was used and the process is outlined in the paragraphs that follow.

The IBI series can be calculated by first finding all peaks in the raw PPG signal and then calculating corresponding IBIs by finding the difference between every two consecutive peaks.

In order to use the Fourier transform, the input signal must consist of evenly spaced samples. This is not the case with the obtained IBI series. We thus interpolate from the original IBI series at a fixed sampling rate in order to construct a new evenly spaced series of samples called the Interpolated IBI series. We then remove the amplitude offset of this series by subtracting the sample mean from each sample and perform a frequency domain analysis on it using the Fast Fourier Transform.

As mentioned in the preceding literature review, the frequency bands of interest include the low frequency and high frequency bands. The low frequency (LF) component contains the total power between 0.04Hz to 0.15Hz, while the high frequency (HF) component contains the total power between 0.15Hz to 0.4Hz. Both of these values can be obtained easily from the DFT performed on the IBI series.

4.2.3 Calculating eye movement entropy

Eye movement entropy is the randomness in a person's eye movement patterns and is calculated in terms of fixation transitions between two or more areas of interest (AOIs). The raw gaze coordinates obtained from an eye-tracking device must first be converted to fixations. Several algorithms may be used to accomplish this, some of the more common ones being based on dispersion-threshold techniques [59], [10], [34].

Once fixations have been recorded, the next task is to determine the AOI in which these fixations occur. Transition frequency matrices for these fixations should then be calculated for each pair of AOIs where $p(i, j)$ denotes the frequency of transitioning from AOI i to

AOI j. The matrices can then be converted to conditional transition probability matrices where each entry (i, j) represents $p(j|i)$, the probability of transitioning to AOI j given a current dwell on AOI i. For each AOI i, the absolute probability $p(i)$ of dwelling on i can be calculated from the percentage dwell time on i.

In order to calculate the eye movement entropy, the following formula may be employed:

$$Entropy = \sum_{i=1}^{n} p(i) \sum_{j=1}^{n} p(j|i) \log_2 p(j|i), i \neq j \tag{5}$$

where n is the number of AOIs, $p(i)$ is the zero-order probability of fixating on AOI i, and $p(j|i)$ is the probability of fixating on AOI j given a current dwell on AOI i.

4.3 Synchronization between Vehicle State and Physiological signals

The sampling rate between each different stream of data differed significantly. The raw PPG and GSR signals were perfectly synchronized (having been jointly measured by the same sensor device) and had a sampling rate of 512 Hz. The eye-tracker, on the other hand, had a sampling rate of 50 Hz, while the vehicle state signals were each sampled (and perfectly synchronized) at 20 Hz.

In addition to the different sampling rates, the physiological signals were also shifted in time with respect to the car signals. This is because the physiological signals were recorded to a separate laptop whose system time differed from that of the PC used on the car. In order to appropriately line up the data, all signal streams first needed to be shifted so that the same timestamp in each stream corresponded to the same timestamp in the other streams. In fact, either the physiological streams or the car streams needed to be shifted, but not both since all of the streams for each group shared the same respective system clocks. The shifting was done through the process of analysing common event points in the physiological and vehicle signal streams. The chosen event points consisted of the time at which the dynamic object was *just* passed, and the time at which the ego vehicle reached the end of the intersection.

For the physiological streams, the eye-tracker video was analyzed in order to manually search for the timestamps of the event points considered. Similarly, the trajectory of the vehicle and its related sensor data were observed on a visualization tool (called "Rviz") in order to search for and record the required timestamps.

In order to find the most accurate clock difference between the two groups of signals, the same event points were recorded for each of the four trials for each participant. The difference in timestamps for corresponding event pairs – in the vehicle state and physiological signals – were then noted for each trial, for a total of 8 differences (2 per trial). The differences were calculated by subtracting the eye-tracker timestamps from the vehicle state timestamps. The 8 differences were averaged over for each participant and the result was added to all timestamps for each physiological signal for that participant. The error in differences was found to be within 500 ms.

Next, in order to combine all signals, a single sampling rate was determined at 20 Hz. Using linear interpolation, all physiological streams were down-sampled to this rate and combined, and the vehicle state signals were interpolated (using nearest-neighbor interpolation) at the resulting physiological timestamps. The resultant dataset, therefore, consisted of all physiological signals aligned and synchronized as accurately as possible with their vehicle state counterparts.

Chapter 5

Analysis

We conducted an initial analysis over all events in order to understand if the events themselves had an effect on the participant's response. Subsequent analyses were then conducted at two levels. First, a coarse-grain analysis was performed at the trial level in order to identify the overall effects of the experimental manipulations during the entire trial. Next, we analysed series of 5 second windows obtained throughout the course of the trial in order to paint a more detailed picture of the effects of each manipulation independent of any specific event.

The next section describes how response and predictor variables were extracted from the complete dataset in order to perform the analysis at each level. This is followed by a section describing the model used to perform the statistical analysis, and a final section describing actual effects observed for each response variable from the participant at each level of analysis.

5.1 Response and Predictor Variables

Response and predictor variables were grouped into two categories: mean and maximum. Mean values for variables were calculated based on the mean absolute values of the response and predictor variables over the interval (i.e., the complete trial, event interval, or the 5 second window) being considered. Similarly, maximum variables were calculated based on the maximum (direction-dependent) values of the variable over the interval.

5.1.1 Response variables

For the event and windowed analysis, maximum-value response variables of interest included:

- **Maximum peak amplitude:** the amplitude of the highest GSR peak found in the interval; if no peak was found, the mean SCL was used.
- **Maximum entropy:** the maximum eye movement entropy over the interval.
- **Maximum heart rate:** the maximum heart rate over the interval.

Mean-value response variables included:

- **Mean SCL:** the mean value of SCL over the interval.
- **Mean entropy:** the mean eye movement entropy over the interval.
- **Mean heart rate:** the mean heart rate over the interval.

In addition to mean and maximum valued response variables, the number of GSR peaks were also studied across all levels of analysis.

The trial-based analysis included all of the response variables above with the additional inclusion of HRV response variables:

- Normalized **low frequency components** of HRV (Lfn)
- Normalized **high frequency components** of HRV (Hfn)
- The ratio **Lfn/Hfn**

5.1.2 Predictor variables

Similar to the response variables, the maximum predictor variables of interest at all levels of analysis included:

- **Maximum longitudinal acceleration (a_{long}^{max}):** the signed maximum value of the longitudinal component of acceleration (positive in the forward direction, and negative in the left).

- **Maximum lateral acceleration (a_{lat}^{max}):** the signed maximum value of the lateral component of acceleration (positive in the right direction, and negative in the left).
- **Maximum longitudinal jerk ($\dot{a}_{long}^{max}$):** the signed maximum value of the longitudinal component of jerk (positive in the forward direction, and negative in the left).
- **Maximum lateral jerk ($\dot{a}_{lat}^{max}$):** the signed maximum value of the lateral component of jerk (positive in the forward direction, and negative in the left).

These signed maximum values depended on the direction of acceleration or jerk, i.e., if the variable was highest in the negative direction, the negative value was considered and not the positive. The reason for this was to preserve directionality in order to distinguish between decelerating to and accelerating from a stop.

Again, the mean-value predictors studied were:

- **Mean longitudinal acceleration ($\bar{a}_{long}$):** the mean absolute value of the longitudinal component of acceleration independent of direction.
- **Mean lateral acceleration ($\bar{a}_{lat}$):** the mean absolute value of the lateral component of acceleration independent of direction.
- **Mean longitudinal jerk ($\bar{\dot{a}}_{long}$):** the mean absolute value of the longitudinal component of jerk independent of direction.
- **Mean lateral jerk ($\bar{\dot{a}}_{lat}$):** the mean absolute value of the lateral component of jerk independent of direction.

The means were calculated independent of direction by taking the absolute value of the variables into account. These values were used to assess the impact of the variable itself without considering its actual direction.

In addition to these predictors, we also considered the presence of a lead and presence of a parked vehicle ($present_{lead}$ and $present_{pass}$, respectively). In cases where these were found to be significant, an additional analysis was performed on a subset of the data consisting of only those samples in which a lead or parked vehicle was present. In these analyses, the **minimum longitudinal distance** d_{long} and **minimum lateral distance** (d_{lat}) to the dynamic object were considered (each of which play a role in car-stop and passing events, respectively).

5.2 Model Description: Linear Mixed Effects

Linear mixed effects (LME) models were used to regress each response variable against predictors. These models are used to account for both fixed effects (i.e. the effects of the predictor variables) as well as random effects, which in our case include the random effects associated with the within subjects nature of the study as well as effects related to the type of event and trial number.

5.2.1 Event models

Simple LME models were used to understand the effect of each event on the response variables. These models simply regressed the response variable on the event type (a factor containing four levels), with the participant ID and trial number as the grouping variable. Two types of orthogonal contrasts were set up for the event factor:

- **Stopping vs. no-stopping events**: these contrasts were used to distinguish intersection and car stopping events from the no-stopping pass and turn events.
- **Car vs. non-car events**: these contrasts were used to distinguish the car events, i.e., passing and car-stop, from the non-car events, turn and intersection stops.

5.2.2 Trial and window level models

At the trial and window level of analysis, the models used differed for different categories of predictors. Mean-valued response variables were regressed on mean-valued predictor variables, while maximum-valued response variables were regressed on maximum-valued predictors. There were also slight differences in structure depending on the level of analysis. In order to study the main effects of each predictor, the following models were used:

- **Trial level:** the random effects initially considered at this level were the participant ID and trial number. However, adding trial number as a random effect did not have any significance and resulted in an inappropriate model structure. The final model, therefore, contained only participant ID as a random effect.

$$
\begin{aligned}
r^{max}_{trial} =& b_0 + b_{pid} + \\
& b_1 d_{long} + b_2 d_{lat} + \\
& b_3 a^{max}_{long} + b_4 dir_{a^{max}_{long}} + b_5 a^{max}_{long} dir_{a^{max}_{long}} + \\
& b_6 a^{max}_{lat} + b_7 dir_{a^{max}_{lat}} + b_8 a^{max}_{lat} dir_{a^{max}_{lat}} + \\
& b_9 \dot{a}^{max}_{long} + b_{10} dir_{\dot{a}^{max}_{long}} + b_{11} \dot{a}^{max}_{long} dir_{\dot{a}^{max}_{long}} + \\
& b_{12} \dot{a}^{max}_{lat} + b_{13} dir_{\dot{a}^{max}_{lat}} + b_{14} \dot{a}^{max}_{lat} dir_{\dot{a}^{max}_{lat}} + \epsilon
\end{aligned}
\tag{5.1}
$$

$$
\begin{aligned}
\bar{r}_{trial} =& b^{'}_0 + b^{'}_{pid} + \\
& b^{'}_1 d_{long} + b^{'}_2 d_{lat} + \\
& b^{'}_3 \bar{a}_{long} + b^{'}_4 \bar{a}_{lat} + b^{'}_5 \bar{\dot{a}}_{long} + b^{'}_6 \bar{\dot{a}}_{lat} + \epsilon^{'}
\end{aligned}
\tag{5.2}
$$

Here, b_0 and $b^{'}_0$ represent the overall intercept terms for the models, b_{pid} and $b^{'}_{pid}$ represent the random intercept terms for participant ID, and ϵ and $\epsilon^{'}$ represent the error.

- **Windowed level:** here, as in the trial level, participant ID and trial number were considered as random effects. In addition, in order to model the t series nature of the subsequent windowed samples from the same participant within a trial, a first-order autoregressive covariance structure was used with the sequence number of the sample as an additional predictor (indicating t).
  Finally, the distance terms were replaced by factor terms representing the presence a lead vehicle ($present_{lead}$) or parked car ($present_{pass}$). This was not necessary at the trial level as the aggregation was performed over the entire trial which necessitates the presence of a lead vehicle and parked car. At the window level, a separate analysis using the distance terms was performed only for the samples in which an agent vehicle was present and only when the presence of the vehicle was significant.

$$
\begin{aligned}
r^{max}_{window} =& b_0 + b_{pid} + b_{trial|pid} + \\
& b_1 t + b_2 present_{lead} + b_3 present_{pass} + \\
& b_4 a^{max}_{long} + b_5 dir_{a^{max}_{long}} + b_6 a^{max}_{long} dir_{a^{max}_{long}} + \\
& b_7 a^{max}_{lat} + b_8 dir_{a^{max}_{lat}} + b_9 a^{max}_{lat} dir_{a^{max}_{lat}} + \\
& b_{10} \dot{a}^{max}_{long} + b_{11} dir_{\dot{a}^{max}_{long}} + b_{12} \dot{a}^{max}_{long} dir_{\dot{a}^{max}_{long}} + \\
& b_{13} \dot{a}^{max}_{lat} + b_{14} dir_{\dot{a}^{max}_{lat}} + b_{15} \dot{a}^{max}_{lat} dir_{\dot{a}^{max}_{lat}} + \epsilon
\end{aligned}
\tag{5.3}
$$

$$\begin{aligned}\bar{r}_{window} =& b_0^{'} + b_{pid}^{'} + b_{trial|pid}^{'} + \\ & b_1 t + b_2^{'} present_{lead} + b_3^{'} present_{pass} + \\ & b_4^{'} \bar{a}_{long} + b_5^{'} \bar{a}_{lat} + b_6^{'} \bar{\dot{a}}_{long} + b_7^{'} \bar{\dot{a}}_{lat} + \epsilon^{'}\end{aligned} \tag{5.4}$$

The meaning of each term is the same as before, except that we now have additional random intercept terms $b_{trial|pid}$ and $b_{trial|pid}^{'}$ that represent the random effect of *trial* nested within the random effect of participant ID.

For the models in which maximum values are considered, the term dir_{x_c} represents a factor term that indicates the direction of the c component of variable x. This term is used to carry out a piece-wise regression on the acceleration (a) and jerk ($\dot{a}$) variables whose signs indicate the direction of the vectors. The $\bar{x}$ terms indicate the mean absolute values of the corresponding variables and are independent of direction. It should be noted that for the trial level, all maximum acceleration and jerk values were in the positive direction and there was no need to consider a factor for direction.

Each term in the model was ordered on the basis of its assumed importance as a predictor. In general, the longitudinal distance from the dynamic object was considered the most important predictor as it indicates the presence of a car up ahead, indicating an upcoming (possibly stressful) stop behind it. The importance of this variable was also corroborated with interview responses from participants, most of whom seemed the most uncomfortable during the car-stop event. Acceleration and jerk terms followed next with jerk appearing after acceleration in order to preserve the order of causality. Furthermore, longitudinal components were assumed to be more important than lateral components based on participant feedback which largely indicated that turns did not cause much discomfort compared to stopping events. This left lateral distance as the last and least important predictor.

Finally, in order to study the interaction effects between variables of interest, the main-effect models were modified to also include a set of interaction terms of interest. We were particularly interested in studying the interaction between longitudinal acceleration (and jerk) and longitudinal distance to the dynamic object. This interaction indicates how longitudinal distance affects participant response for different values of acceleration and jerk. For maximum-valued models the terms capturing these interactions were:

1. $d_{long} \cdot a_{long}^{max} \cdot dir_{a_{long}^{max}}$ for **acceleration**, and $d_{long} \cdot \dot{a}_{long}^{max} \cdot dir_{\dot{a}_{long}}$ for **jerk** at the trial level.

2. $present_{lead} \cdot a_{long}^{max} \cdot dir_{a_{long}^{max}}$ for **acceleration**, and $present_{lead} \cdot \dot{a}_{long}^{max} \cdot dir_{\dot{a}_{long}}$ for **jerk** at the window level.

For mean-valued models, the corresponding interaction terms were:

1. $d_{long} \cdot \bar{a}_{long}$ for **acceleration**, and $d_{long} \cdot \bar{\dot{a}}_{long}$ for **jerk** and the trial level.
2. $present_{lead} \cdot \bar{a}_{long}$ for **acceleration**, and $present_{lead} \cdot \bar{\dot{a}}_{long}$ for **jerk** at the window level.

After adding in all the expected fixed effects based on our study rationale, a few models contained additional interaction terms that were found to be significant, even if these terms were not originally expected. These additional interaction terms are highlighted later in the appropriate section on the response variable.

Significance levels considered

For all analyses, a significance level of $p < 0.05$ was considered as the threshold for significance. A weak significance was considered at $p < 0.1$.

Table 5.1: The effect of Stopping vs. Non-stopping events on GSR response variables.

	Dependent variable:		
	Max. Pk. Amp.	Num. Pks.	Mean SCL
	b (se)	b (se)	b (se)
EventStopvsNoStop	0.157*** (0.018)	0.496*** (0.068)	0.066*** (0.010)
EventTurnvsPass	−0.048* (0.025)	−0.216** (0.097)	0.003 (0.014)
EventIntervsCarStop	−0.081*** (0.025)	−0.057 (0.097)	−0.043*** (0.014)
Constant	0.214*** (0.042)	1.127*** (0.132)	0.038*** (0.010)
Observations	314	314	314
Log Likelihood	−98.374	−516.993	109.739
Akaike Inf. Crit.	210.749	1,047.987	−205.477
Bayesian Inf. Crit.	236.995	1,074.233	−179.232
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5.2: The effect of Stopping vs. Non-stopping events on Maximum Heart Rate and Maximum Entropy response variables.

	Dependent variable:	
	Max. HR	Max. Entropy
	b (se)	b (se)
EventStopvsNoStop	3.952^{***} (1.188)	0.005^{**} (0.002)
EventTurnvsPass	-2.796^{*} (1.679)	0.004 (0.003)
EventIntervsCarStop	0.857 (1.683)	-0.001 (0.003)
Constant	125.966^{***} (3.756)	0.248^{***} (0.009)
Observations	303	239
Log Likelihood	$-1,378.283$	417.954
Akaike Inf. Crit.	2,770.567	-821.907
Bayesian Inf. Crit.	2,796.563	-797.572
Note:		*p<0.1; **p<0.05; ***p<0.01

5.3 Understanding the Role of Events

Before conducting a more detailed analysis, an analysis was first conducted on the events employed in order to understand the high level influence of each event type. This analysis was carried out irrespective of the driving style parameters, and used planned contrasts as well as post-hoc Tukey's HSD (Honestly Significant Difference) tests in order to more clearly distinguish the contribution of each event to the overall level of response. All response variables were aggregated over the duration of each entire event (non-event regions were not considered).

5.3.1 Planned contrasts

Figures 5.1 and 5.2 depict the variation in response with event type for maximum peak amplitude, number of peaks, maximum heart rate, maximum entropy, mean SCL, mean heart rate, and mean entropy response variables. A quick glance at each of the plots tells us that stopping events had an overall larger impact on response than non-stopping events, with car-stops having a slightly higher impact than intersection stops. Furthermore, the mean heart rate did not seem to be affected by any event in particular, although the maximum heart rate did show significant differences for different events.

Tables 5.1 and 5.2 confirm our overall interpretation of the plots: the "Stop vs. No-

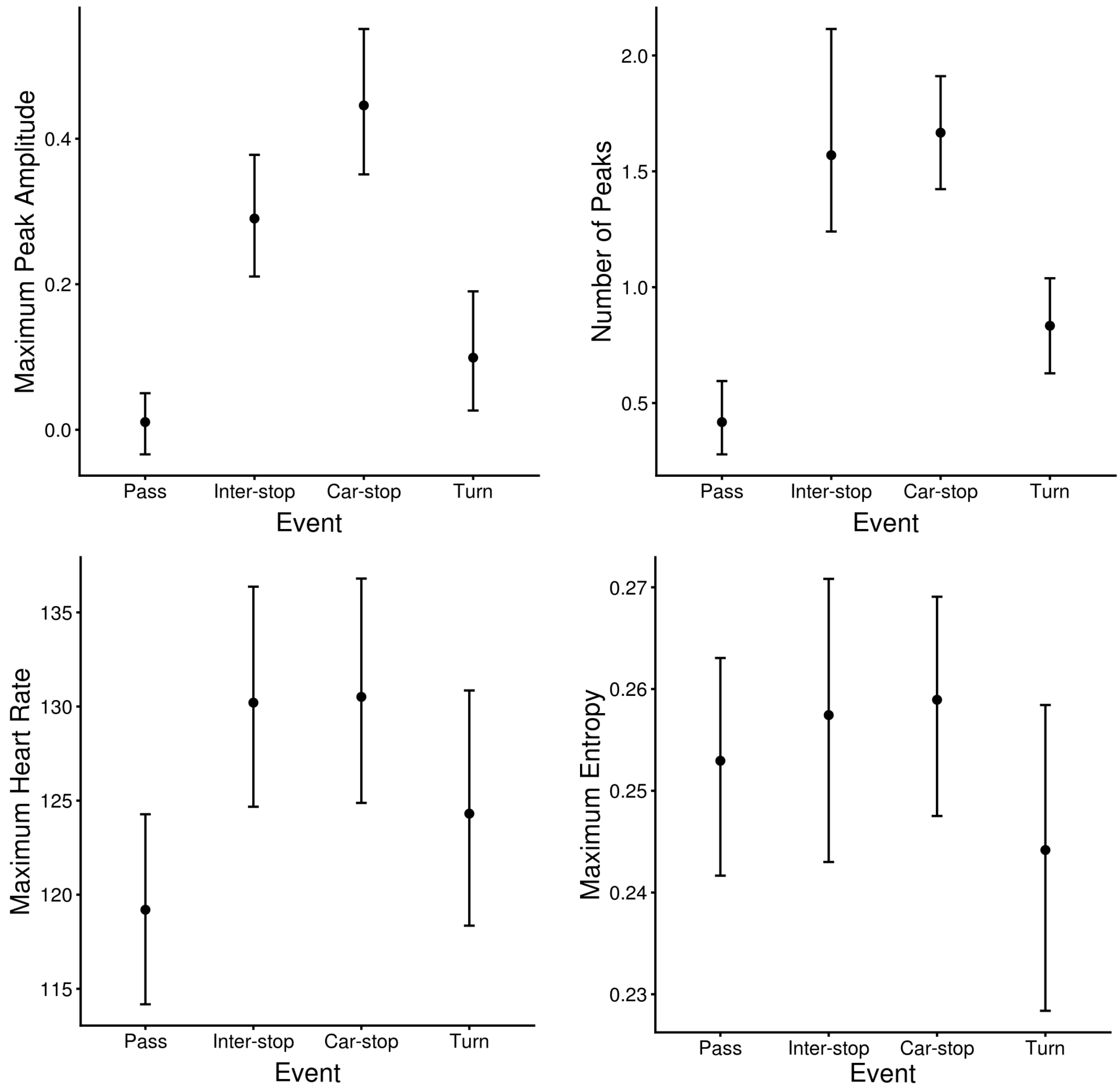


Figure 5.1: The effect of event type on Maximum Peak Amplitude, Number of Peaks, Maximum Heart Rate, and Maximum Entropy.

stop" contrast has the highest significance level for all maximum-value response variables as well as for mean SCL. Mean heart rate and mean entropy, however, did not show any significant difference between stopping and non-stopping events and so their tables were

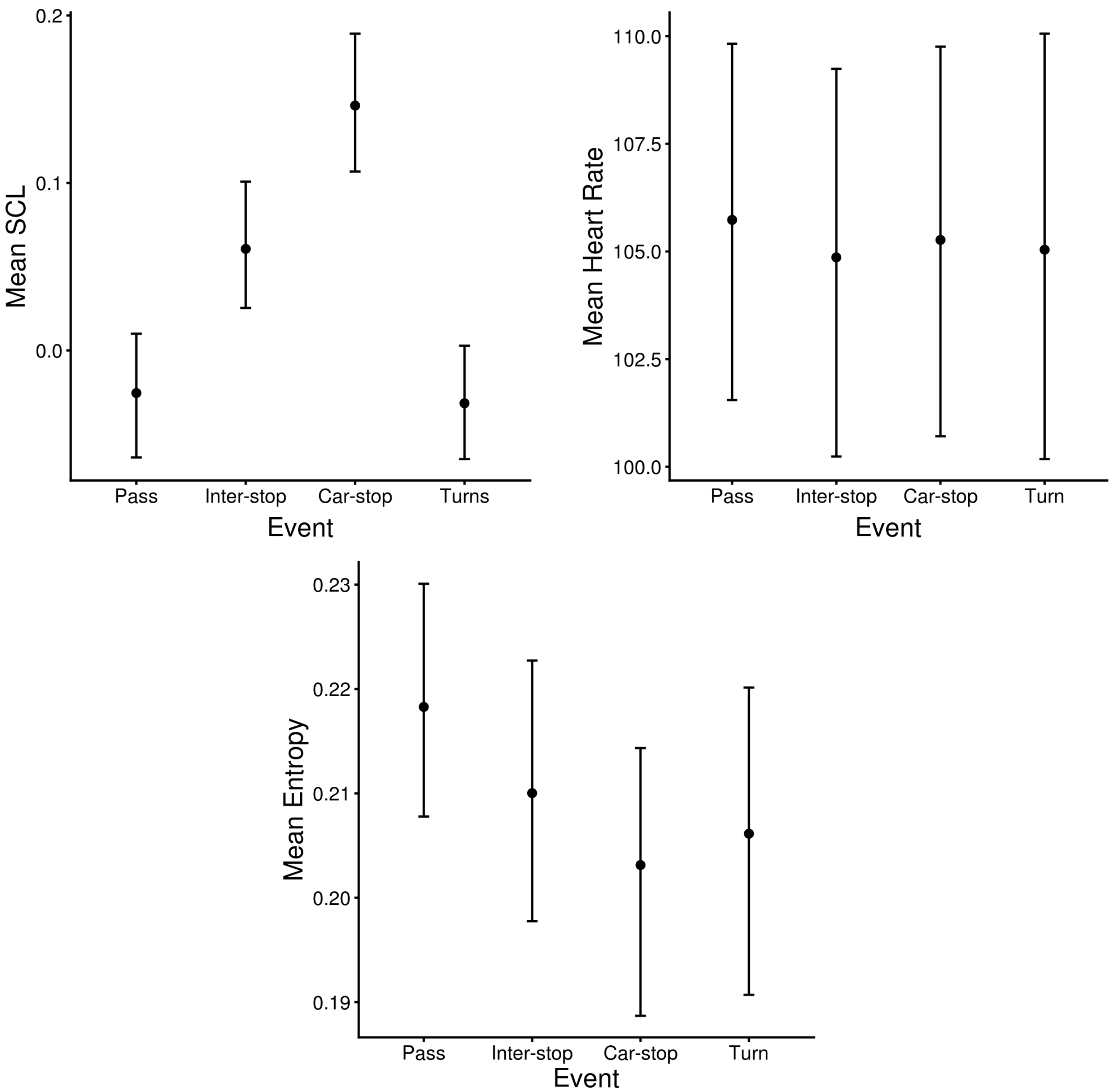


Figure 5.2: The effect of event type on Mean SCL, Mean Heart Rate, and Mean Entropy.

not included here.

Within the category of stopping events, there were significant differences between stopping at an intersection and stopping behind a car for maximum peak amplitude and mean

SCL, but not for the maximum entropy response variable. This can be verified from the plots in Figures 5.1 and 5.2, where the Car-stop event has a response higher than that for all other event types for these variables. For the non-stopping events, significant differences in the number of peaks existed between turning and passing a car. The same effect was also seen with a weak significance for maximum peak amplitude and maximum heart rate. In general, turning events generate a higher response than passing a car.

5.3.2 Post-hoc tests

In order to gain further insight into the differences between individual pairs of events, post-hoc tests using Tukey's HSD were used to compare pair-wise differences in means between the four events. As with the planned contrasts, these tests were carried out for each response variable. Once again, the mean heart rate and mean entropy did not show any significant differences in means between any pair of events.

GSR variables

All three GSR variables showed significant differences between most pairs of events.

There were significant differences in maximum peak amplitude between passing and stopping at an intersection ($d = 0.28, p < 0.001$), passing and stopping behind a car ($d = 0.44, p < 0.001$), stopping behind a car and intersection ($d = 0.16, p < 0.01$), turning and stopping at an intersection ($d = -0.18, p < 0.01$), and turning and stopping behind a car ($d = -0.34, p < 0.001$).

With respect to number of peaks, there were significant differences between passing and both stopping events: intersection-stops ($d = 1.15, p < 0.001$) and car-stops($d = 1.27, p < 0.001$). Turns were also significantly different from intersection-stops ($d = -0.71, p < 0.01$) and car-stops ($d = -0.83, p < 0.001$).

Finally, for mean SCL, intersection and car stops were significantly different from passing ($d = 0.09, p < 0.01$ for intersection stops, and $d = 0.17, p < 0.001$ for car stops) as well as turning ($d = -0.09, p < 0.01$ for intersection stops, and $d = -0.17, p < 0.001$ for car stops) events. The type of stop, intersection as opposed to a car stop, was also significantly different ($d = 0.08, p < 0.01$).

Heart rate variables

There was no significant difference between individual events for the mean heart while the maximum heart rate was significantly different only between stopping events and passing ($d = 11.07, p < 0.01$ for intersection stops, and $d = 11.33, p < 0.01$ for stopping behind a car).

Entropy variables

Unlike the other response variables, both maximum and mean entropy did not show any significant differences between the individual events even though, for maximum entropy, the Stop vs. No-stop contrast was significant.

5.3.3 Takeaways: the role of events

On the whole, there was a clear difference of effect between stopping and no-stopping events. Within the stopping events, the presence of a car elicited a generally higher skin response, while within the no-stopping events, turn elicited a generally higher skin response as well as increased the maximum heart rate.

5.4 Using LME and Linear Regression to Estimate Predictors

In this chapter we report all effects found for all response variables at the **window** level of analysis. For the trial level, however, we only report the effects from models that had significance in at least one predictor variable. An exhaustive list of effects for all variables at the trial level can be found in the Appendices.

5.4.1 GSR

GSR variables of interest included the number of peaks, the maximum peak amplitude, and the mean SCL. Significant main and interaction effects were found for all three response variables at the window level, but only for number of peaks at the trial level. Tables of effects are provided for all three variables at the window level but only for the number of peaks at the trial level.

Table 5.3: Main effects for GSR variables at the window level.

	Dependent variable:		
	Max. Pk. Amp.	Num. Pks.	Mean SCL
	b (se)	b (se)	b (se)
t	−0.002 (0.002)	−0.001 (0.004)	0.00005 (0.002)
$present_{lead}$	0.121*** (0.037)	0.432*** (0.082)	0.067*** (0.022)
a_{long}^{max}	−0.026 (0.019)	−0.084* (0.043)	
$dir_{a_{long}^{max}}$	−0.077*** (0.027)	−0.094* (0.055)	
a_{lat}^{max}	0.036 (0.100)	−0.002 (0.222)	
$dir_{a_{lat}^{max}}$	−0.029 (0.037)	−0.028 (0.082)	
$\dot{a}_{long}^{max}$	−0.025 (0.022)	−0.066 (0.048)	
$dir_{\dot{a}_{long}^{max}}$	−0.013 (0.015)	−0.022 (0.039)	
$\dot{a}_{lat}^{max}$	−0.024 (0.036)	−0.104 (0.087)	
$dir_{\dot{a}_{lat}^{max}}$	0.039** (0.019)	0.029 (0.045)	
$\bar{a}_{long}$			0.023 (0.026)
$\bar{a}_{lat}$			−0.002 (0.017)
$\bar{\dot{a}}_{long}$			0.147** (0.059)
$\bar{\dot{a}}_{lat}$			−0.087* (0.048)
$present_{pass}$	−0.033 (0.029)	−0.041 (0.059)	−0.033 (0.021)
$a_{long}^{max}{:}dir_{a_{long}^{max}}$	0.114*** (0.035)	0.330*** (0.076)	
$a_{lat}^{max}{:}dir_{a_{lat}^{max}}$	−0.024 (0.101)	0.018 (0.225)	
$\dot{a}_{long}^{max}{:}dir_{\dot{a}_{long}^{max}}$	0.111*** (0.038)	0.333*** (0.085)	
$\dot{a}_{lat}^{max}{:}dir_{\dot{a}_{lat}^{max}}$	−0.033 (0.063)	0.056 (0.147)	
Constant	0.009 (0.049)	0.183 (0.115)	−0.043** (0.020)
Observations	1,700	1,700	1,700
Log Likelihood	−208.368	−1,580.307	298.015
Akaike Inf. Crit.	456.736	3,200.614	−572.030
Bayesian Inf. Crit.	565.504	3,309.382	−506.769

Note: *p<0.1; **p<0.05; ***p<0.01

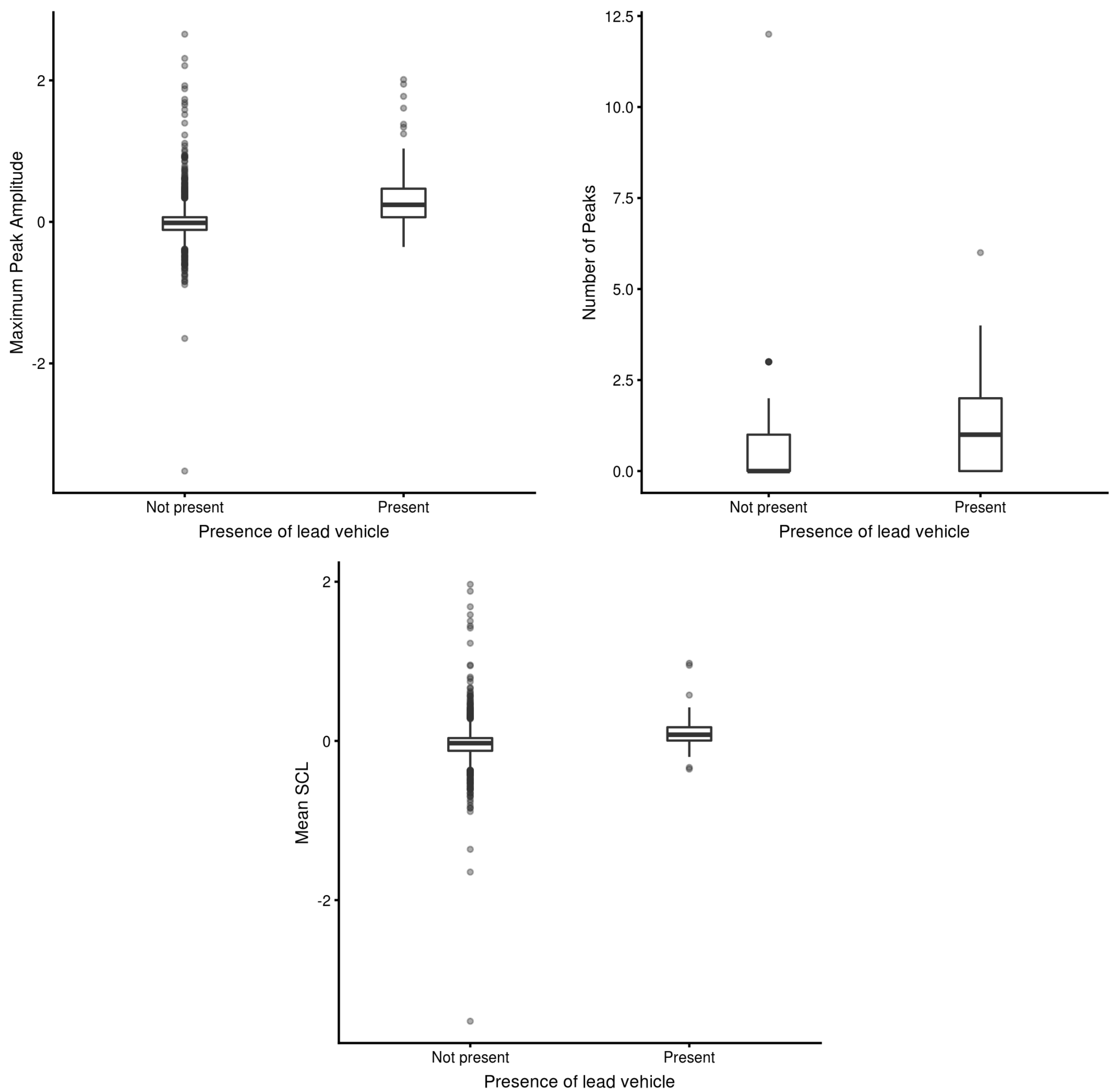


Figure 5.3: Influence of lead vehicle presence on GSR variables at the window level.

Random effects

At the window level, there was no significant variance in intercepts across participants or across trials for the maximum peak amplitude or the mean SCL. For the number of peaks,

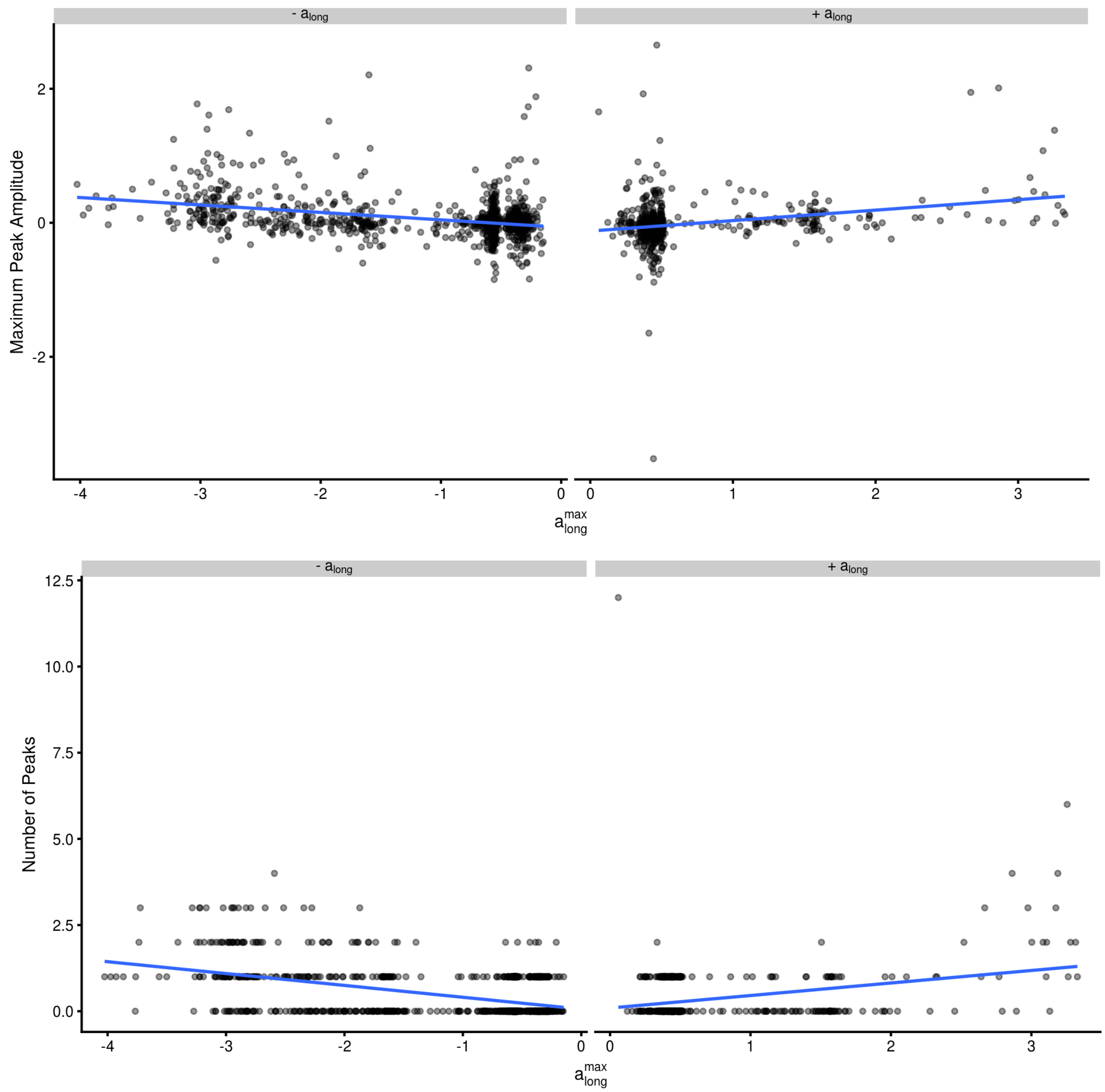
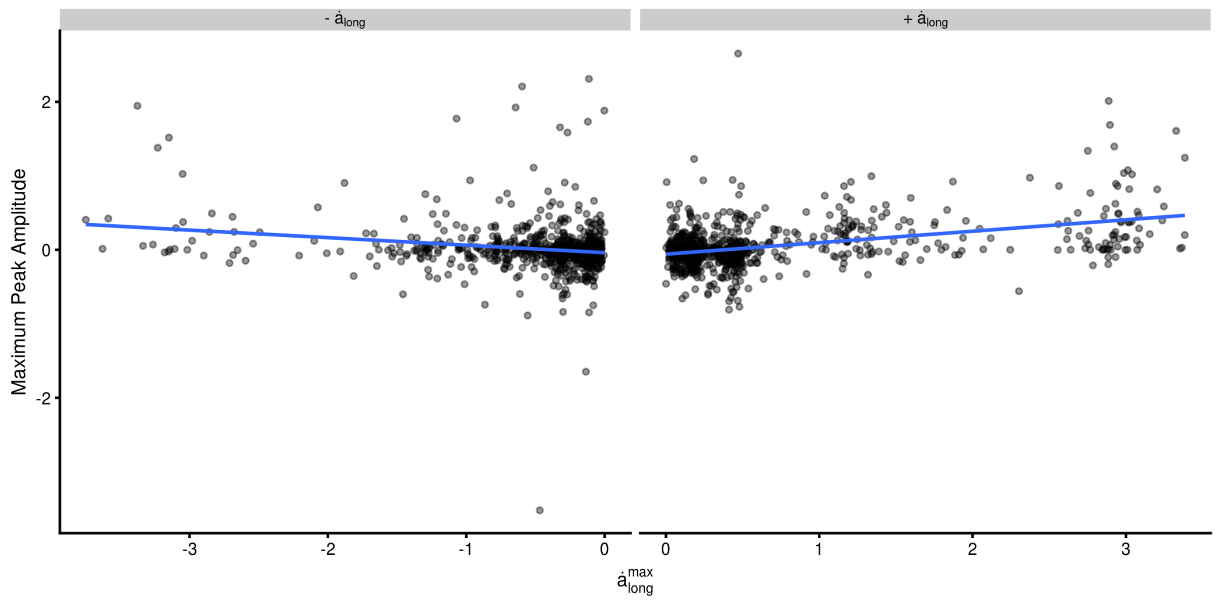


Figure 5.4: Influence of a_{long}^{max} on Max. Peak Amplitude and No. of Peaks at the window level.



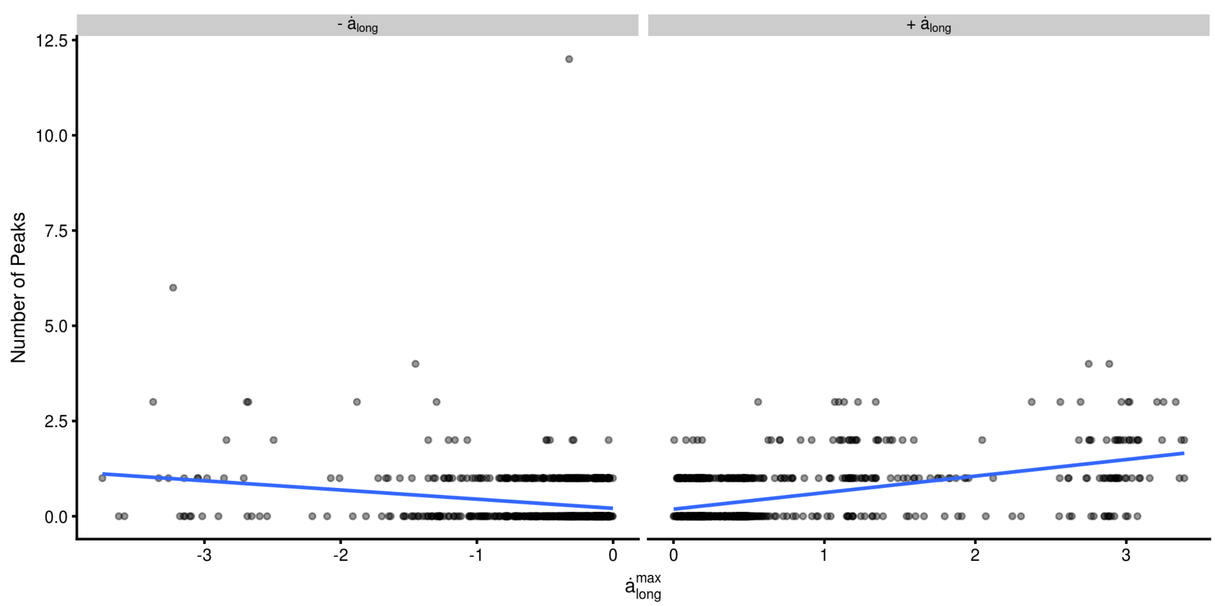


Figure 5.5: Influence of $\dot{a}_{long}^{max}$ on Max. Peak Amplitude and No. of Peaks at the window level.

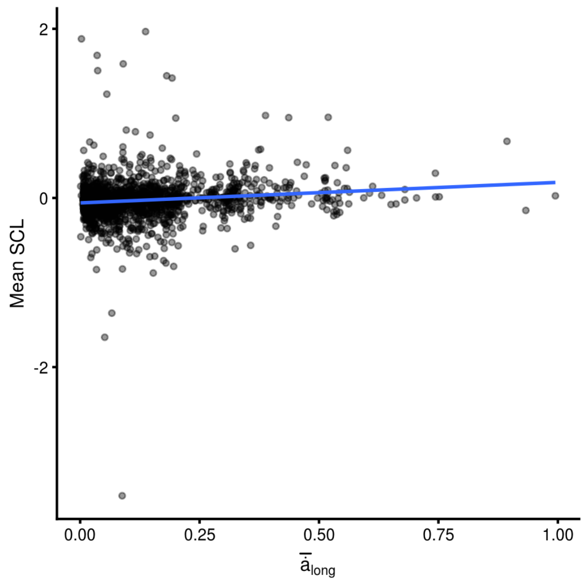


Figure 5.6: Influence of $\bar{a}_{long}$ on Mean SCL at the window level.

however, there was significant variation in intercepts across participants, SD = 0.176 (95% CI: 0.123, 0.252), $\chi^2(1) = 94.33, p < 0.0001$, but not across trials.

Finally, adding an autoregressive covariance structure significantly improved the model in the case of mean SCL, $\chi^2(1) = 516.45, p < 0.0001$ as well as the maximum peak amplitude, $\chi^2(1) = 205.98, p < 0.0001$, but did not result in any improvement for the number of peaks response.

At the trial level, only the model for number of peaks was considered as other models did not have any significant main or interaction effects. For this model, there was significant variance in intercepts across participants, SD = 3.75 (95% CI: 2.67, 5.26), $\chi^2(1) = 57.73, p < 0.0001$.

Main effects

Table 5.3 lists the overall main effects at the window level of all predictor variables for different GSR response variables. The presence of a lead vehicle ($present_{lead}$) was a significant predictor for all three response variables, i.e., the maximum peak amplitude (Max. Pk. Amp.), the number of peaks (Num. Pks.), as well as the mean SCL (Mean SCL).

Table 5.4: Trial level main and interaction effects for Number of Peaks. The first column represents the model for the main effects, while interaction effects are listed in the second column.

	Dependent variable:	
	Num. Pks.	
	b (se)	b (se)
d_{long}	0.036^{**} (0.017)	0.024 (0.073)
a_{long}^{max}	0.654 (0.638)	-4.410 (2.669)
a_{lat}^{max}	0.328 (0.516)	-3.708^{**} (1.787)
$\dot{a}_{long}^{max}$	-0.159 (1.379)	-0.184 (2.214)
$\dot{a}_{lat}^{max}$	-0.105 (2.229)	-1.073 (6.104)
d_{lat}	-0.015 (0.014)	-0.015 (0.014)
d_{long}:a_{long}^{max}		-0.009 (0.070)
d_{long}:$\dot{a}_{long}^{max}$		0.003 (0.055)
a_{long}^{max}:a_{lat}^{max}		1.740^{**} (0.751)
$\dot{a}_{long}^{max}$:$\dot{a}_{lat}^{max}$		0.329 (1.950)
Constant	6.610^{**} (2.866)	18.812^{**} (8.288)
Observations	80	80
Log Likelihood	-200.661	-196.958
Akaike Inf. Crit.	419.323	419.917
Bayesian Inf. Crit.	440.761	450.883
Note:		*p<0.1; **p<0.05; ***p<0.01

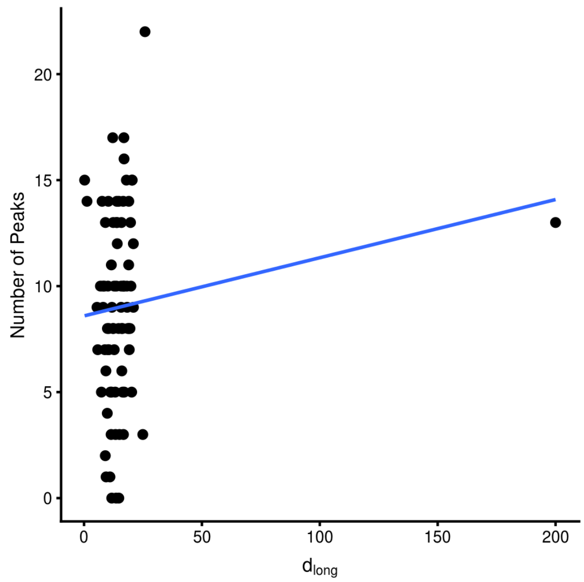


Figure 5.7: Relationship between d_{long} and Number of Peaks at the window level of analysis. The data-points consist of the subset of samples for which the lead vehicle was present.

The maximum signed longitudinal acceleration and jerk variables (a and $\dot{a}$) were significant predictors for both the maximum peak amplitude and number of peaks, while mean longitudinal jerk ($\bar{\dot{a}}$) significantly predicted the mean SCL.

Figure 5.3 depicts the relationship between each GSR response variable with the presence of a lead vehicle at the window level. The box-plots indicate that the median for each response variable is consistently higher in the condition when the lead vehicle is present, suggesting that the presence of a car results in a higher magnitude of skin response.

At the window level, figures 5.4 and 5.5 shows the relationship between the maximum peak amplitude and the number of peaks with a_{long}^{max} and $\dot{a}_{long}^{max}$ respectively while figure 5.6 shows the influence of $\bar{\dot{a}}_{long}$ on the mean SCL. In all cases, the magnitude of each predictor has a positive relationship with the response variable, regardless of its direction.

At the trial level, significant main effects were found only for the number of peaks response variable. Table 5.4 lists all effects for this response variable. Of these, only d_{long} was found to be significant. However, as seen in Figure 5.7 this relationship is contrary to what was expected and, by extension, what the window level analysis tells us.

Table 5.5: Interaction effects of interest for GSR variables at the window level.

	Dependent variable:		
	Max. Pk. Amp.	Num. Pks.	Mean SCL
	b (se)	b (se)	b (se)
$present_{lead}1{:}a^{max}_{long}{:}dir_{a^{max}_{long}}$	0.249** (0.097)	1.134*** (0.214)	
$present_{lead}1{:}\dot{a}^{max}_{long}{:}dir_{\dot{a}^{max}_{long}}$	0.122* (0.071)	0.504*** (0.159)	
$present_{lead}1{:}\bar{a}_{long}$			0.198** (0.078)
$present_{lead}1{:}\dot{\bar{a}}_{long}$			−0.279** (0.126)
Constant	0.032 (0.050)	0.258** (0.115)	−0.041** (0.020)
Observations	1,700	1,700	1,700
Log Likelihood	−197.583	−1,546.397	310.801
Akaike Inf. Crit.	447.167	3,144.794	−589.603
Bayesian Inf. Crit.	588.565	3,286.192	−502.589
Note:			*p<0.1; **p<0.05; ***p<0.01

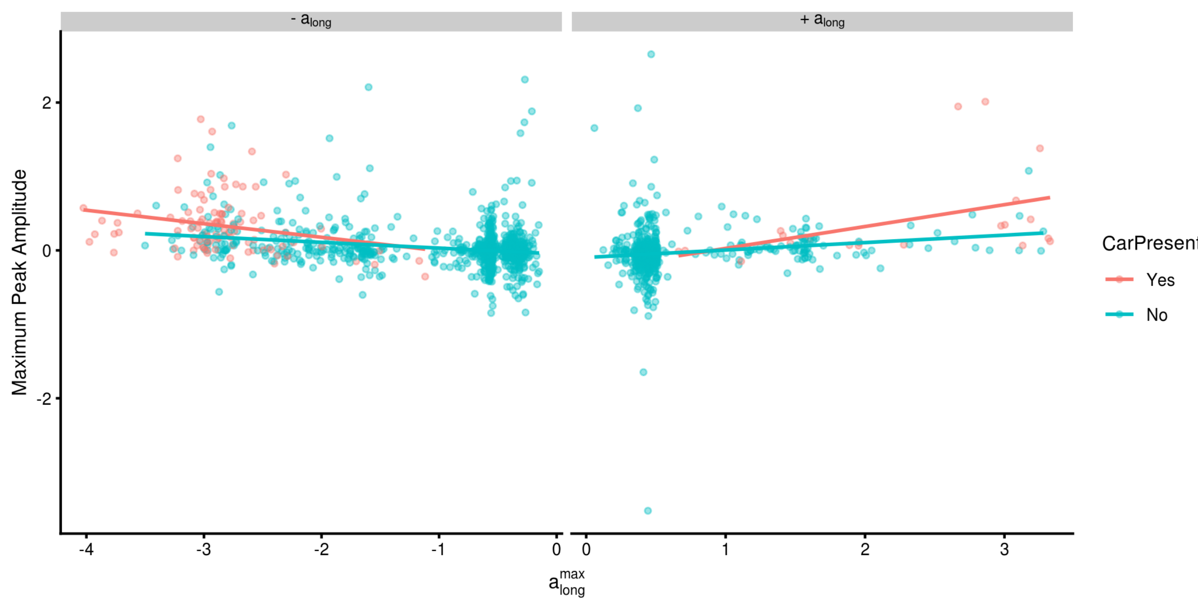


Figure 5.8: Interaction effects of $present_{lead}$ with a^{max}_{long} for Maximum Peak Amplitude.

Interaction effects

From Table 5.5 we observe significant interaction effects at the window level between $present_{lead}$ and a^{max}_{long} for maximum peak amplitude and number of peaks, as well as between

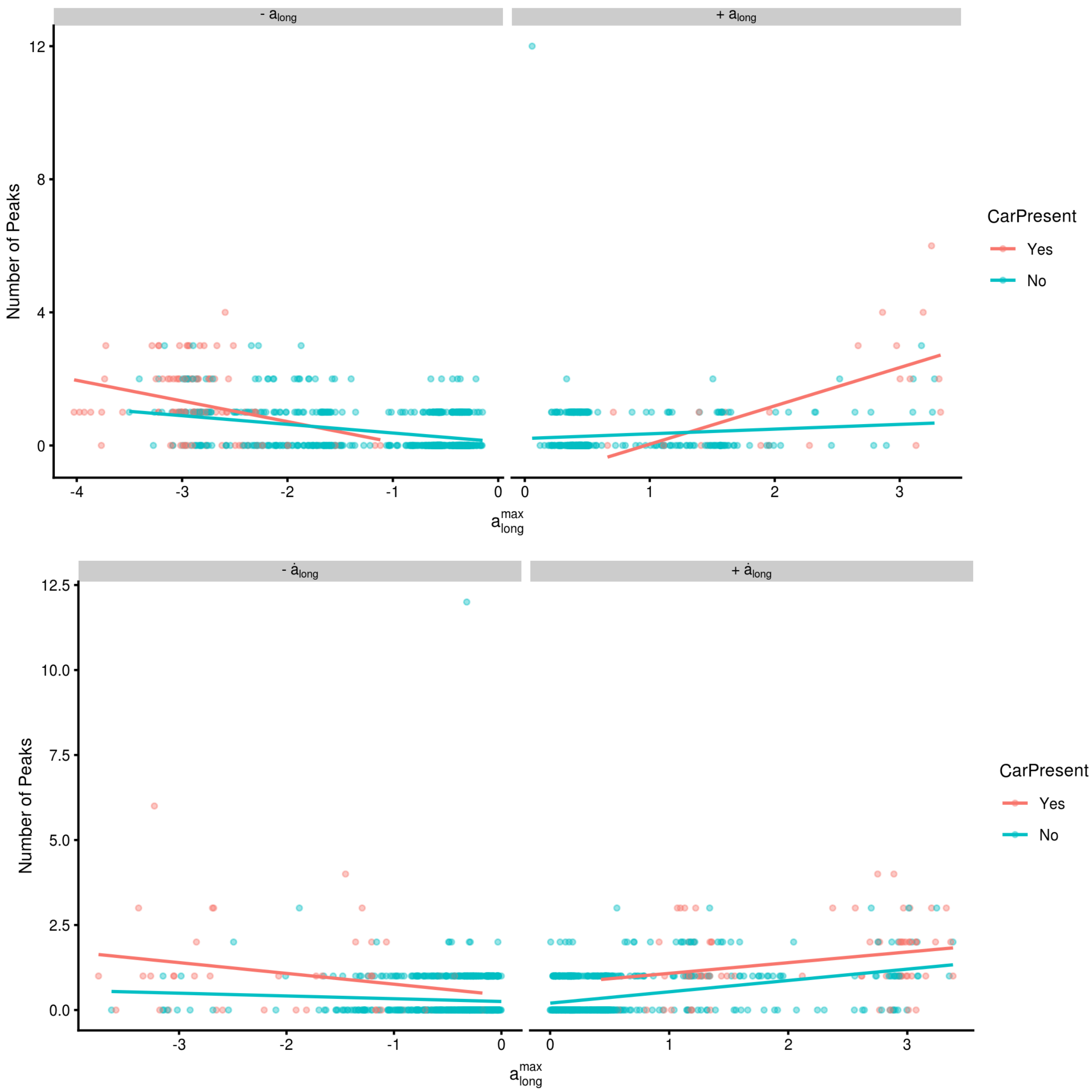


Figure 5.9: Interaction effects of $present_{lead}$ with a_{long}^{max}, and $\dot{a}_{long}^{max}$ for Number of Peaks.

$present_{lead}$ and $\dot{a}_{long}^{max}$ for the number of peaks response variable. Likewise, there are similar interaction effects for mean SCL between $present_{lead}$ and $\bar{a}_{long}$, and between $present_{lead}$ and $\bar{\dot{a}}_{long}$.

Figures 5.8 and 5.9 show clear interaction effects between $present_{lead}$ and the a_{long}^{max} (for maximum peak amplitude and number of peaks) and $\dot{a}_{long}^{max}$ variables (for number of peaks). In all cases, the presence of a lead vehicle results in a greater as well as sharper response in the number of peaks and maximum peak amplitude, and this response is higher for higher values of acceleration and jerk independent of direction.

There were no interaction effects observed at the trial level for any GSR response variables.

Takeaways

Not only does the presence of a lead vehicle have a significant overall main effect for all skin response variables, but it also modulates the effect of longitudinal acceleration and jerk for all of these variables. These variables also have significant effects on their own for all GSR response variables.

Although we are able to see an unusual relationship between d_{long} and the number of peaks at the trial level, we believe that this relationship might be due to the fact that not much variation in distance would have been observed at this level of analysis in order to form a more complete picture. This is because the distance recorded would have been the minimum over each complete trial, which might not provide a large enough variation in order to see an effect.

Conversely, as all three GSR variables (irrespective of whether they were maximum or mean valued) were significantly affected by the same predictor variables, the analysis shows that GSR response, in general, is highly susceptible to changes in vehicle state.

5.4.2 HR and HRV

HR and HRV variables of interest included the maximum and mean heart rate, Lfn, Hfn and the ratio Lfn/Hfn. Due to the short intervals considered at the window level, responses of Lfn, Hfn, and the ratio between them were not taken into consideration at this level of analysis. Since each trial was approximately 3 minutes in length, these HRV variables were analysed at this level.

Significant main and interaction effects were found for all HR variables at the window level. However, there were significant main (but no interaction effects) found only for mean heart rate at the trial level. HRV variables showed no significant main or interaction effects for this level of analysis.

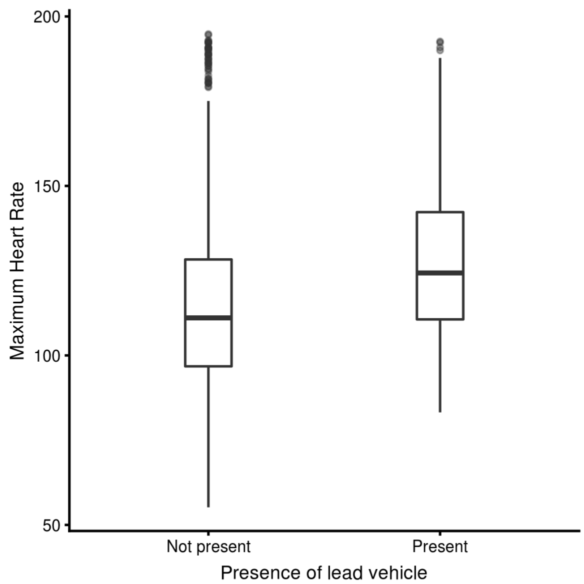


Figure 5.10: Influence of $present_{lead}$ on Max. HR at the window level.

Random effects

At the window level, there was significant variance in intercepts across participants, SD = 14.72 (95% CI: 10.45, 20.73), $\chi^2(1) = 631.92, p < 0.0001$, as well as across trials, SD = 6.02 (95% CI: 3.84, 9.44), $\chi^2(1) = 149.57, p < 0.0001$, for the maximum HR response. The mean HR response also showed significant variance in intercepts across participants, SD= 14.16 (95% CI: 10.11, 19.83), $\chi^2(1) = 785.03, p < 0.0001$, and trials, SD= 5.33 (95% CI: 3.41, 8.33), $\chi^2(1) = 161.58, p < 0.0001$.

Finally, adding an autoregressive covariance structure significantly improved both models: in the case of maximum HR, $\chi^2(1) = 433.92, p < 0.0001$, while for mean HR, $\chi^2(1) = 494.98, p < 0.0001$.

At the trial level, the model for mean heart showed significant variance in intercepts across participants, SD = 12.08 (95% CI: 8.59, 16.99), $\chi^2(1) = 51.76, p < 0.0001$.

Table 5.6: Main effects for HR variables at the window level.

	Dependent variable:	
	Max. HR	Mean HR
	b (se)	b (se)
t	−0.013 (0.138)	−0.125 (0.115)
$present_{lead}$	6.553*** (2.163)	0.731 (1.478)
a^{max}_{long}	−1.954* (1.150)	
$dir_{a^{max}_{long}}$	−2.950* (1.628)	
a^{max}_{lat}	−13.713** (5.895)	
$dir_{a^{max}_{lat}}$	3.877* (2.205)	
$\dot{a}^{max}_{long}$	−0.762 (1.317)	
$dir_{\dot{a}^{max}_{long}}$	0.593 (0.879)	
$\dot{a}^{max}_{lat}$	3.653 (2.324)	
$dir_{\dot{a}^{max}_{lat}}$	−0.310 (1.058)	
$\bar{a}_{long}$		0.946 (1.837)
$\bar{a}_{lat}$		1.442 (1.196)
$\bar{\dot{a}}_{long}$		−4.476 (4.220)
$\bar{\dot{a}}_{lat}$		−3.170 (3.375)
$present_{pass}$	−1.345 (1.764)	−1.162 (1.435)
$a^{max}_{long}{:}dir_{a^{max}_{long}}$	5.554*** (2.119)	
$a^{max}_{lat}{:}dir_{a^{max}_{lat}}$	15.503** (6.033)	
$\dot{a}^{max}_{long}{:}dir_{\dot{a}^{max}_{long}}$	3.354 (2.380)	
$\dot{a}^{max}_{lat}{:}dir_{\dot{a}^{max}_{lat}}$	−7.821* (4.191)	
Constant	109.982*** (4.535)	106.595*** (3.516)
Observations	1,700	1,700
Log Likelihood	−7,221.909	−6,939.275
Akaike Inf. Crit.	14,483.820	13,902.550
Bayesian Inf. Crit.	14,592.590	13,967.810
Note:		*p<0.1; **p<0.05; ***p<0.01

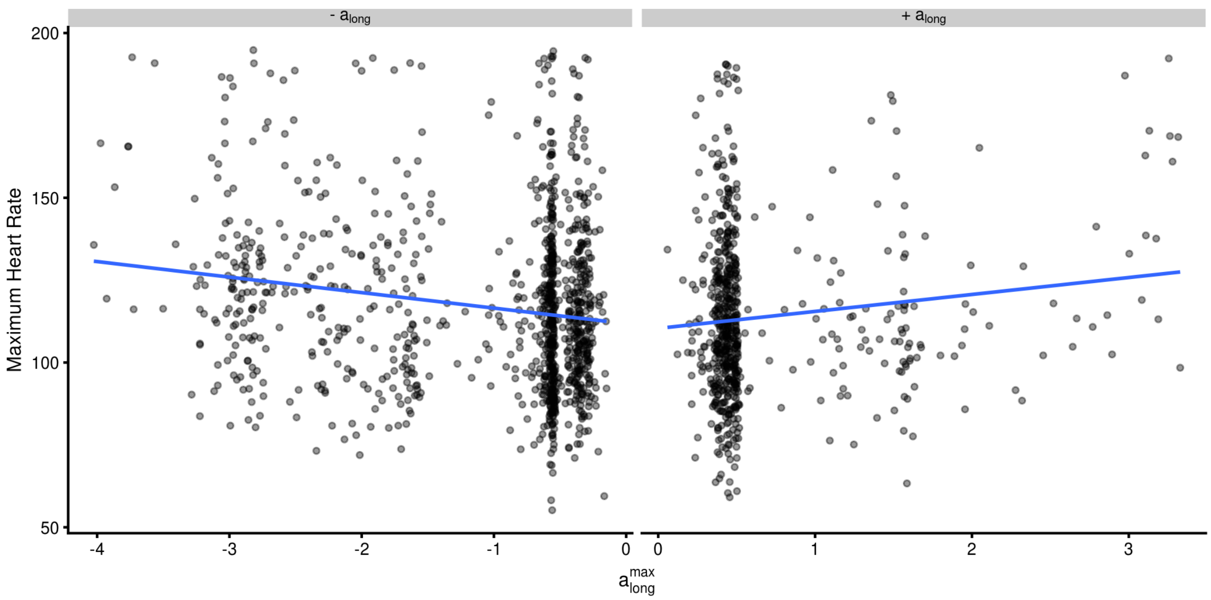


Figure 5.11: Influence of a_{long}^{max} on Max. HR at the window level.

Table 5.7: Main effects for Mean HR at the trial level.

	Dependent variable:
	MeanHr
LongDist	−0.060 (0.118)
MeanLongAcc	−15.234 (40.399)
MeanLatAcc	135.740** (65.692)
MeanLongJerk	−5.557 (62.026)
MeanLatJerk	−152.855 (102.145)
LatDist	−0.031 (0.053)
Constant	84.063*** (24.751)
Observations	80
Log Likelihood	−298.819
Akaike Inf. Crit.	615.638
Bayesian Inf. Crit.	637.076
Note:	*p<0.1; **p<0.05; ***p<0.01

Main effects

Table 5.6 lists the overall main effects for HR variables, i.e., the maximum HR (Max. HR) and mean HR (Mean HR). In this case, $present_{lead}$, a_{long}^{max}, and a_{lat}^{max} were significant

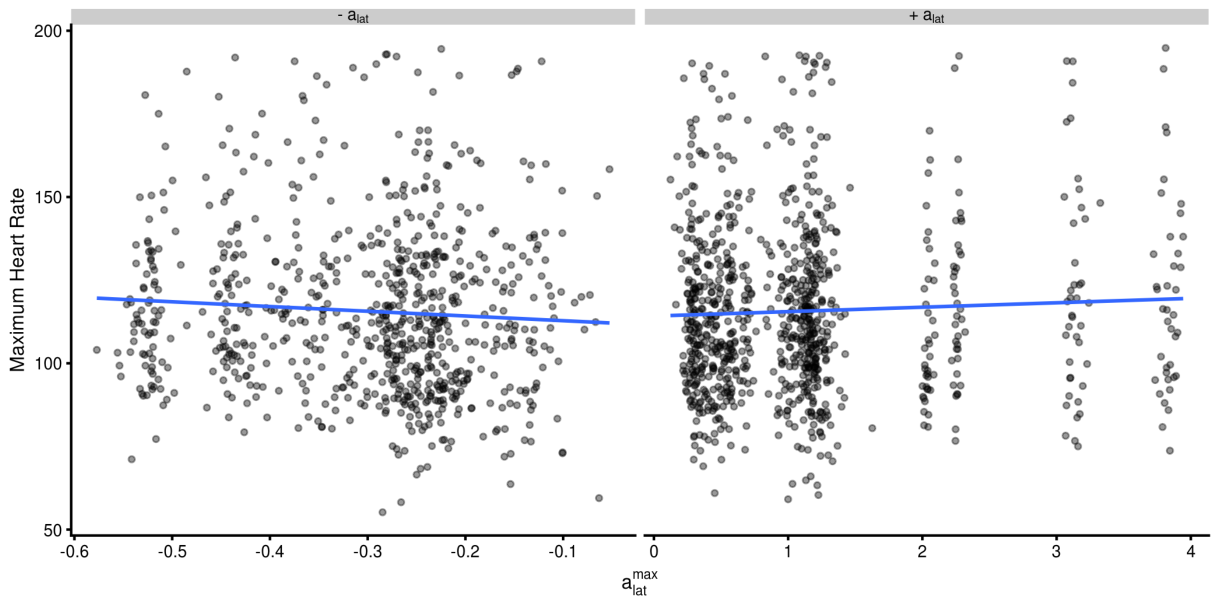


Figure 5.12: Influence of a_{lat}^{max} on Max. HR at the window level.

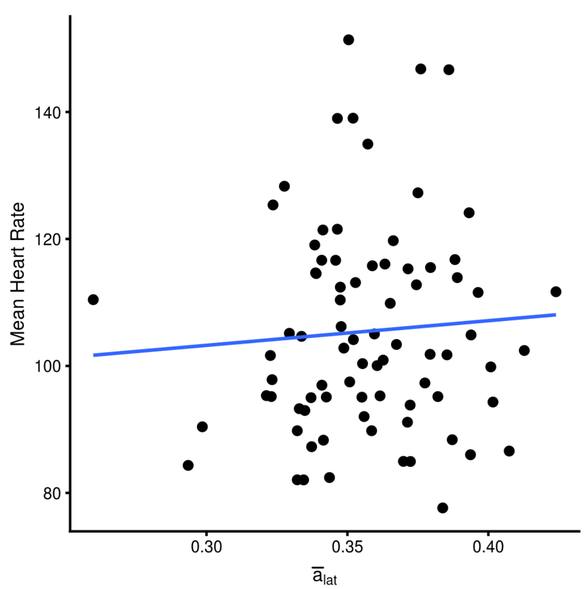


Figure 5.13: Influence of $\bar{a}_{lat}$ on Mean HR at the trial level.

predictors, but only of the maximum HR; there was no significant predictor for mean HR.

As with the GSR response variables, Figure 5.10 depicts the relationship between $present_{lead}$ with the maximum HR. As before, the maximum heart increases within the presence of a vehicle. Positive relationships are seen in Figure 5.11 with the magnitude of a^{max}_{long}, in both positive and negative direction. In addition to the longitudinal component of acceleration, the lateral component also shows a positive relationship (of the magnitude) with maximum HR in either direction (Figure 5.12).

At the trial level, Table 5.7 lists the main effects found for mean heart rate. Only $\bar{a}_{lat}$ had a significant and positive effect as seen in Figure 5.13. This result was consistent with what was seen in the window level analysis.

Table 5.8: Interaction effects of interest for HR variables at the window level.

	Dependent variable:	
	Max. HR	Mean HR
	b (*se*)	*b* (*se*)
$present_{lead}1{:}a^{max}_{long}{:}dir_{a^{max}_{long}}$	6.783 (5.625)	
$present_{lead}1{:}\dot{a}^{max}_{long}{:}dir_{\dot{a}^{max}_{long}}$	9.118** (4.120)	
$present_{lead}1{:}\bar{a}_{long}$		−2.959 (5.380)
$present_{lead}1{:}\dot{\bar{a}}_{long}$		9.781 (8.654)
Constant	109.774*** (4.551)	106.644*** (3.518)
Observations	1,700	1,700
Log Likelihood	−7,215.367	−6,936.485
Akaike Inf. Crit.	14,482.730	13,904.970
Bayesian Inf. Crit.	14,624.130	13,991.980
Note:	*p<0.1; **p<0.05; ***p<0.01	

Interaction effects

At the window level, the only interaction effect is between $present_{lead}$ and $\dot{a}^{max}_{long}$ for the maximum heart rate response variable (Table 5.8). Figure 5.14 depicts this interaction effect: it can be clearly seen that the presence of a car results in an overall higher heart rate. However, it is not clear why this heart rate decreases as the magnitude of jerk increases in both directions.

There were no significant interaction effects for any response variable at the trial level.

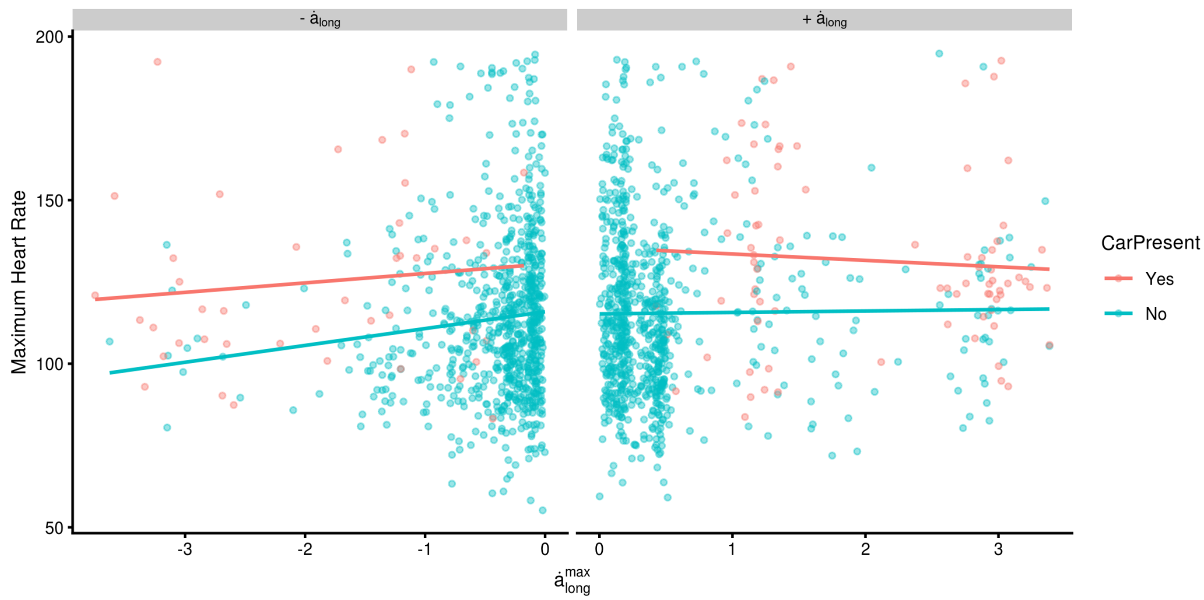


Figure 5.14: Interaction effect between $present_{lead}$ and $\dot{a}_{long}^{max}$ for Maximum HR

Takeaways

Overall, the effects on heart rate were similar to those seen with the GSR response variables. There were overall main effects of lead vehicle presence, and longitudinal acceleration, with an additional affect of lateral acceleration.

The additional affect of lateral acceleration at both the window and trial levels suggests that heart rate, in particular, might be more susceptible to turns than other response variables. This point is further bolstered by additional evidence from Table 5.6, which indicates a weak effect of lateral jerk for the maximum heart rate response.

What is a bit unusual, however, is the interaction effect between the longitudinal jerk and presence of a lead vehicle: while there is a clear effect on heart rate in the presence of a vehicle irrespective of the value of jerk, within the category of presence or absence, the expected increase in heart rate with an increase in jerk magnitude does not occur. This may simply be due to the fact that longitudinal jerk itself does not really play much of a role in affecting the response as can be seen from the main effect table.

Finally, unlike the GSR response variables, the mean heart rate does not seem to be too responsive to changes in vehicle state - the only effect seen on this variable was at the trial level which, in itself, is not the most informative model when compared to the lower resolution window level model. Nevertheless, given the significant impact of lateral acceleration on heart rate (and not skin response), this response variable could suitably supplement skin response in a system that is used to gauge comfort or stress from both components of acceleration and jerk.

5.4.3 Eye movement entropy

Eye movement entropy variables of interest included the maximum and mean entropy.

At the window level of analysis, significant main as well as interaction effects were found for both entropy response variables. At the trial level, there was a significant interactione effect only for maximum entropy.

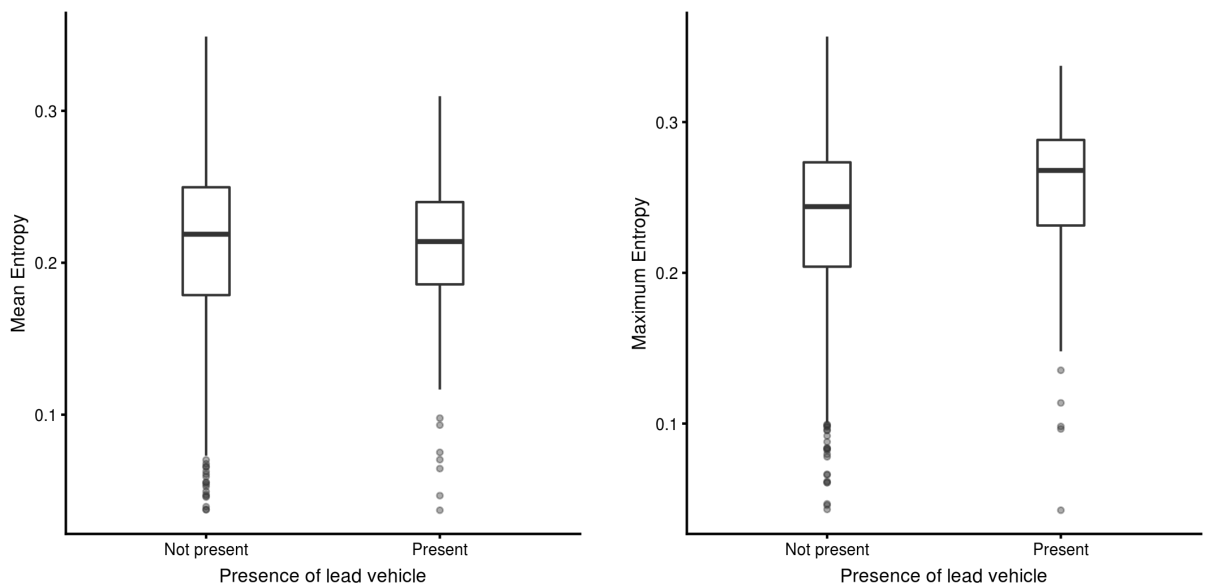


Figure 5.15: Influence of $present_{lead}$ on Max. and Mean Entropy at the window level.

Random effects

At the window level, once again, the random effects and autoregressive co-variance structure played significant roles in shaping both models. However, despite the apparent significance of the co-variance structure, the model itself resulted in an infinite upper

Table 5.9: Main effects for eye movement entropy variables at the window level.

	Dependent variable:	
	Max. Entropy	Mean Entropy
	b (se)	b (se)
t	−0.0004 (0.0003)	−0.0003 (0.0003)
$present_{lead}$	0.014** (0.006)	−0.004 (0.005)
a^{max}_{long}	0.002 (0.003)	
$dir_{a^{max}_{long}}$	−0.004 (0.004)	
a^{max}_{lat}	−0.014 (0.017)	
$dir_{a^{max}_{lat}}$	0.011* (0.006)	
$\dot{a}^{max}_{long}$	−0.004 (0.004)	
$dir_{\dot{a}^{max}_{long}}$	0.001 (0.003)	
$\dot{a}^{max}_{lat}$	−0.003 (0.007)	
$dir_{\dot{a}^{max}_{lat}}$	0.0003 (0.003)	
$\bar{a}_{long}$		−0.003 (0.007)
$\bar{a}_{lat}$		0.002 (0.004)
$\bar{\dot{a}}_{long}$		0.012 (0.014)
$\bar{\dot{a}}_{lat}$		0.014 (0.012)
$present_{pass}$	0.002 (0.005)	0.004 (0.004)
$a^{max}_{long}{:}dir_{a^{max}_{long}}$	0.002 (0.006)	
$a^{max}_{lat}{:}dir_{a^{max}_{lat}}$	0.013 (0.017)	
$\dot{a}^{max}_{long}{:}dir_{\dot{a}^{max}_{long}}$	0.009 (0.007)	
$\dot{a}^{max}_{lat}{:}dir_{\dot{a}^{max}_{lat}}$	0.011 (0.012)	
Constant	0.223*** (0.011)	0.206*** (0.008)
Observations	1,352	1,352
Log Likelihood	2,338.157	2,292.127
Akaike Inf. Crit.	−4,638.315	−4,562.255
Bayesian Inf. Crit.	−4,539.337	−4,504.952
Note:		*p<0.1; **p<0.05; ***p<0.01

confidence level indicating that it was unsuitable. Hence, the co-variance structure was dropped and the default was used for both entropy response variables. For maximum entropy, intercepts varied significantly for participants, SD = 0.03 (95% CI: 0.02, 0.045), $\chi^2(1) = 456.00, p < 0.0001$, as well as across trials, SD = 0.016 (95% CI: 0.012, 0.022), $\chi^2(1) = 82.50, p < 0.0001$. Likewise mean entropy also showed significant variance in intercepts across participants, SD= 0.032 (95% CI: 0.022,0.047), $\chi^2(1) = 461.35, p < 0.0001$, and trials, SD= 0.017 (95% CI: 0.013, 0.022), $\chi^2(1) = 80.36, p < 0.0001$.

Main effects

Table 5.9 lists the overall main effects for eye movement entropy variables, .i.e., the maximum and mean entropy (Max. Entropy and Mean Entropy, respectively). In addition to the vehicle predictor variable $present_{lead}$, t was found to be a significant predictor of maximum entropy.

As expected, the presence of a lead vehicle raises the maximum entropy as shown in (Figure 5.15); unexpected, however, is the opposite relationship of $present_{lead}$ with mean entropy as seen in the same figure. Here, the vehicle's presence actually seems to decrease the overall mean entropy.

Table 5.10: Interaction effects of interest for eye movement entropy variables at the window level.

	Dependent variable:	
	Max. Entropy	Mean Entropy
	b (se)	b (se)
$present_{lead}1{:}a^{max}_{long}{:}dir_{a^{max}_{long}}$	0.032* (0.017)	
$present_{lead}1{:}\dot{a}^{max}_{long}{:}dir_{\dot{a}^{max}_{long}}$	−0.008 (0.013)	
$present_{lead}1{:}\bar{a}_{long}$		0.025 (0.020)
$present_{lead}1{:}\dot{\bar{a}}_{long}$		0.024 (0.032)
Constant	0.227*** (0.011)	0.207*** (0.008)
Observations	1,352	1,352
Log Likelihood	2,179.457	2,277.199
Akaike Inf. Crit.	−4,270.914	−4,522.398
Bayesian Inf. Crit.	−4,041.703	−4,439.049
Note:		*p<0.1; **p<0.05; ***p<0.01

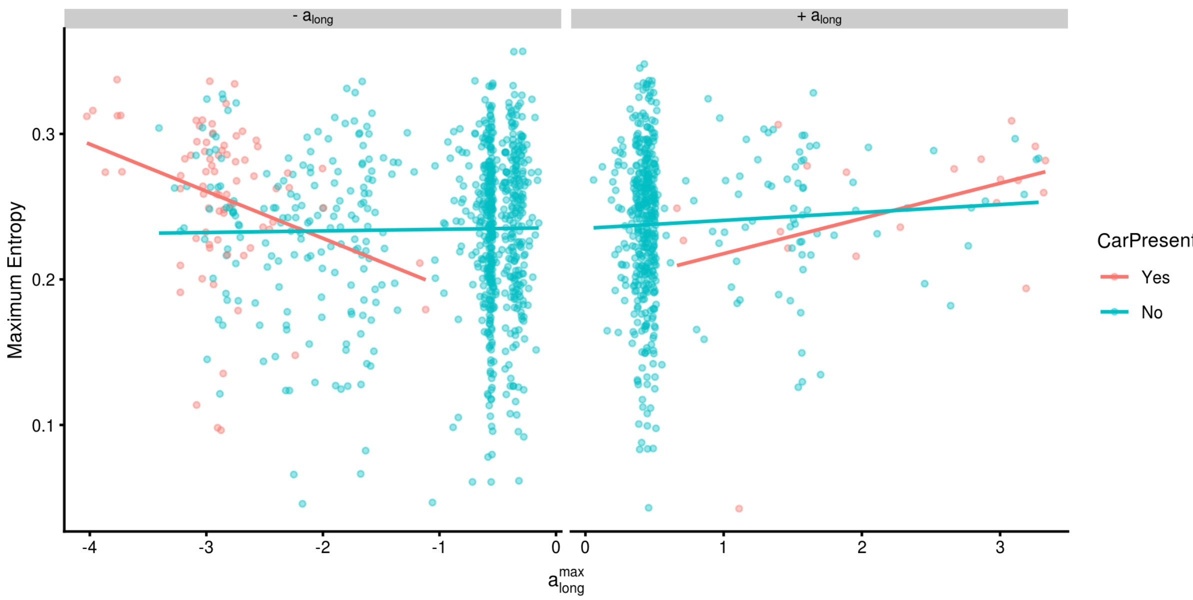


Figure 5.16: Interaction effects of $present_{lead}$ with a_{long}^{max} for Maximum Entropy.

Table 5.11: Interaction effects for Maximum Entropy at the trial level

	Dependent variable:
	Max. Entropy
d_{long}	-0.008^{***} (0.003)
a_{long}^{max}	-0.105^{**} (0.046)
a_{lat}^{max}	-0.009 (0.022)
$\dot{a}_{long}^{max}$	0.154^{**} (0.074)
$\dot{a}_{lat}^{max}$	-0.137 (0.098)
d_{lat}	-0.007 (0.007)
d_{long}:a_{long}^{max}	0.008^{***} (0.003)
d_{long}:$\dot{a}_{long}^{max}$	-0.002 (0.002)
Constant	0.926^{***} (0.144)
Observations	66
Log Likelihood	46.844
Akaike Inf. Crit.	-71.687
Bayesian Inf. Crit.	-47.601
Note:	*p<0.1; **p<0.05; ***p<0.01

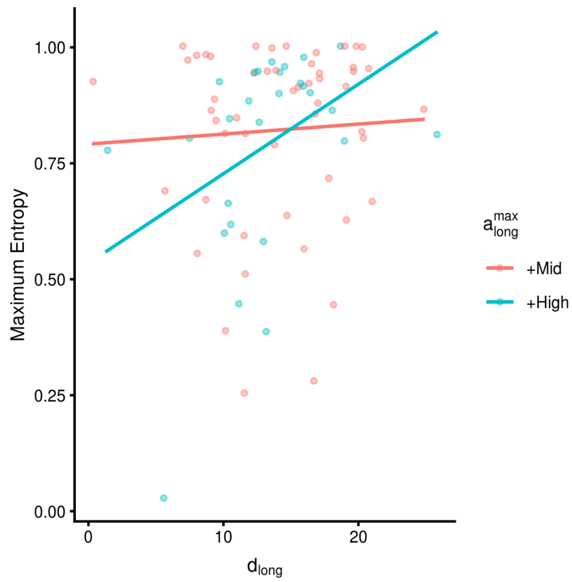


Figure 5.17: Interaction effects of d_{long} with a_{long}^{max} for Maximum Entropy at the trial level.

Interaction effects

Table 5.10 lists the interaction effects observed for eye movement entropy variables. A significant interaction occurs between $present_{lead}$ and a_{long}^{max} for maximum entropy. Figure 5.16 shows the interaction effect in this context. There was a sharper increase in the maximum entropy when a car was present. Moreover, the magnitude of the response was greater for higher magnitudes of a_{long}^{max}, when the car was present.

No significant interaction effects were observed for mean entropy.

At the trial level, a significant interaction effect was observed between d_{long} and a_{long}^{max} as seen in Table 5.11. As seen in Figure 5.17, the entropy actually decreases the closer the ego vehicle comes to the lead vehicle.

Takeaways

Except for the presence of a lead vehicle, there does not appear to be any main effect from any of the other predictors possibly indicating that the entropy is solely affected by the presence or absence of a vehicle rather than its actual motion. However, there does appear to be an interaction effect between the presence of the lead vehicle and a_{long}^{max} which is consistent with what we have seen before with the other response variables at the window level. The effect is not consistent with other results at the trial level, but a possible reason for this could be the low level of information contained at this level of analysis.

5.5 Analysing the interaction effect: zooming in on the stopping distance

Previously, we have looked at the interaction affect between $present_{lead}$ and a_{long}^{max} as well as $\dot{a}_{long}^{max}$. It is quite clear the the presence of a lead vehicle has an overall influence on the effect of longitudinal acceleration and jerk for nearly all of the response variables considered.

However, we hypothesized that the actual distance of the ego vehicle from the agent vehicle itself should also play a role in generating a response in the participant. In order to test this hypothesis, separate models were constructed that focused only on the portion of the trial that had the lead vehicle present. Significant effects were found only in the model for maximum heart rate, while a weak effect ($p < 0.1$) was found for the number of peaks response. For the sake of demonstration, both models are considered in this analysis (Table 5.12).

The interaction effect is clear for the number of peaks response: higher magnitudes of acceleration in either direction result in an overall higher response, which is all the more higher when the ego vehicle is closer to the lead vehicle, i.e., when d_{long} is lower. A similar effect is seen for maximum heart rate when the acceleration is positive. In this case, the expected behavior is seen when the acceleration is high, whereas for lower acceleration values the effect is exactly the opposite with the slopes for the two cases intersecting.

5.6 The Direction of Acceleration and Jerk

In order to determine whether positive or negative acceleration (as well as jerk) had a greater effect, the marginal effects of each acceleration and jerk component were also computed for each response variable. These were computed only for the maximum-valued variables since these were only variables that took direction into account. Furthermore, as the analysis in the trial level was not able to capture direction even in the maximum valued predictors (since the interval was too big), these effects were computed only at the window level. The marginal effects for each response variable are listed in Table 5.13. Terms for which at least one component was significant are displayed in bold print. Expected results

Table 5.12: Interaction effects for the Number of Peaks and Maximum Heart Rate response variables for situations in which a lead vehicle is present.

	Dependent variable:	
	Num. Pks.	Max. HR
	b (se)	b (se)
t	−0.032 (0.039)	−0.434 (1.090)
d_{long}	0.041** (0.016)	1.492*** (0.442)
a^{max}_{long}	−0.933** (0.364)	−16.658 (10.038)
$dir_{a^{max}_{long}}$	0.905 (1.917)	−12.191 (52.056)
a^{max}_{lat}	−0.794 (2.088)	−6.628 (57.510)
$dir_{a^{max}_{lat}}$	−0.433 (0.568)	−1.321 (15.652)
$\dot{a}^{max}_{long}$	−0.316 (0.310)	6.910 (8.527)
$dir_{\dot{a}^{max}_{long}}$	0.889 (0.667)	2.359 (17.958)
$\dot{a}^{max}_{lat}$	−2.443*** (0.605)	−0.898 (16.607)
$dir_{\dot{a}^{max}_{lat}}$	0.425 (0.314)	−4.081 (8.510)
a^{max}_{long}:$dir_{a^{max}_{long}}$	2.112** (0.816)	40.929* (21.989)
a^{max}_{lat}:$dir_{a^{max}_{lat}}$	0.528 (2.207)	15.800 (60.083)
$\dot{a}^{max}_{long}$:$dir_{\dot{a}^{max}_{long}}$	0.387 (0.479)	−9.400 (12.942)
$\dot{a}^{max}_{lat}$:$dir_{\dot{a}^{max}_{lat}}$	3.994*** (1.144)	−7.083 (31.257)
d_{long}:a^{max}_{long}	0.012 (0.011)	0.613* (0.298)
d_{long}:$dir_{a^{max}_{long}}$	−0.008 (0.022)	−1.020 (0.597)
d_{long}:$\dot{a}^{max}_{long}$	0.016* (0.008)	−0.316 (0.227)
d_{long}:$dir_{\dot{a}^{max}_{long}}$	−0.014 (0.011)	0.077 (0.287)
d_{long}:a^{max}_{long}:$dir_{a^{max}_{long}}$	−0.030* (0.016)	−0.957** (0.446)
d_{long}:$\dot{a}^{max}_{long}$:$dir_{\dot{a}^{max}_{long}}$	−0.021 (0.015)	0.462 (0.409)
Constant	−2.447** (1.045)	88.364*** (28.473)
Observations	121	121
Log Likelihood	−86.844	−486.500
Akaike Inf. Crit.	223.688	1,023.000
Bayesian Inf. Crit.	293.583	1,092.895
Note:		*p<0.1; **p<0.05; ***p<0.01

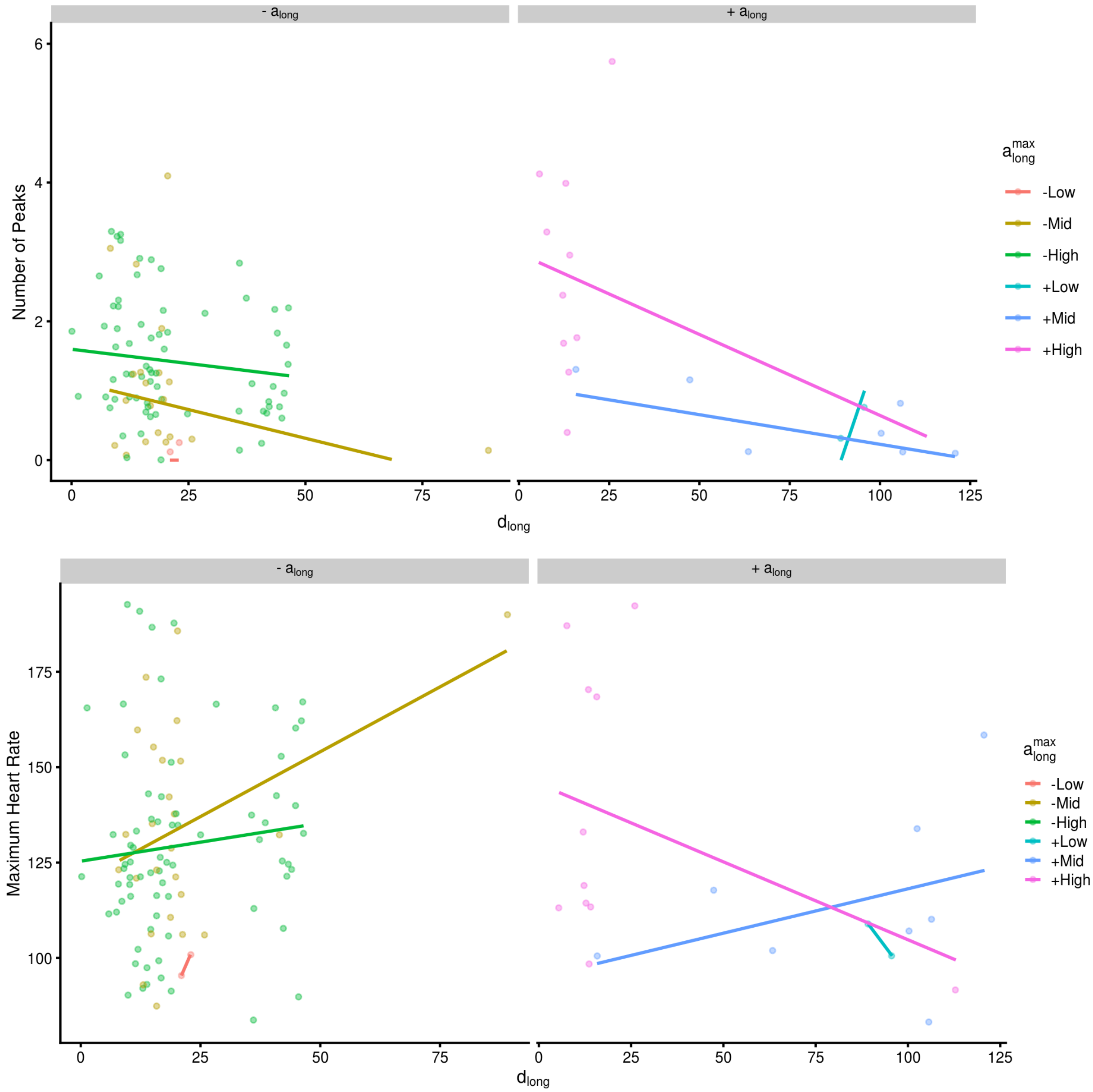


Figure 5.18: Interaction between d_{long} and a_{long}^{max} for Number of Peaks and Maximum Heart Rate response variables.

Table 5.13: Marginal effects of acceleration and jerk components on each response variable.

Response	Predictors							
	a_{long}^{max}		$\dot{a}_{long}^{max}$		a_{lat}^{max}		$\dot{a}_{lat}^{max}$	
	-ve	+ve	-ve	+ve	-ve	+ve	-ve	+ve
Max. Pk. Amp.	**-0.026**	**0.088**	**-0.024**	**0.086**	0.036	0.011	-0.024	-0.058
Num. Pks.	**-0.084**	**0.246**	**-0.066**	**0.267**	-0.002	0.016	-0.104	-0.048
Max. HR	**-1.954**	**3.599**	-0.762	2.592	**-13.713**	**1.789**	3.653	-4.168

would indicate a negative slope in the negative direction and a positive slope in the positive direction.

The significant terms listed in the table agree with the expected results and can also be verified from earlier plots. In order to understand the effect size, we simply need to consider the magnitude irrespective of direction. For all cases in which longitudinal acceleration and jerk are significant, the positive direction of acceleration has a greater magnitude of effect than the negative direction. For lateral acceleration, which is significant only for the maximum heart rate response, it is the negative direction of acceleration that has the larger effect size.

The results suggest that for the longitudinal component of acceleration and jerk, the positive direction of a_{long}^{max} and $\dot{a}_{long}^{max}$ has a larger impact on physiological response than does the negative direction. On the other hand, lateral acceleration shows the opposite effect: the negative direction has a larger impact on heart rate than the positive direction. This effect is surprising because left turns (negative direction) were not really present during driving. This may then be indicative of a confounding variable present during turns, or perhaps just noisy readings from the accelerometer.

5.7 Analysing Self-report Scores

We analysed the relationship between physiological response and on-the-fly self-reported comfort and anxiety scores, as well as between physiological response and CSAI-2 scores. On-the-fly scores were on a Likert scale with integer scores renging from 1 to 10, while the somatic sub-scale of the CSAI-2 questionnaire was used with integer scores ranging from 4 to 36. Since th on-the-fly scores were recorded after each event, these scores were analyzed at an event level, in which physiological responses were aggregated over event intervals. The CSAI-2 were necessarily analyzed at the trial level as these were measured once at the end of each trial. As before, separate models were constructed for maximum-valued responses and mean-values responses, with the number of peaks appearing in all models as this is neither maximum nor mean-valued variable.

Tables 5.14 and 5.15 show the effects of each physiological response on the on-the-fly and CSAI-2 scores respectively. For both on-the-fly anxiety and comfort as well as CSAI-2 scores, number of peaks is a significant predictor, while the mean heart rate is also a significant predictor for comfort (but not for comfort or anxiety).

Table 5.14: Effects of physiological response on self-reported on-the-fly scores.

	Dependent variable:			
	Anxiety		Comfort	
	(1)	(2)	(3)	(4)
NumPks	0.118 (0.072)	0.140** (0.064)	−0.227*** (0.074)	−0.253*** (0.066)
MaxPkAmp	0.329 (0.263)		−0.415 (0.274)	
MaxHr	0.0005 (0.004)		−0.001 (0.004)	
MeanScl		0.666 (0.459)		−0.883* (0.486)
MeanHr		−0.008 (0.006)		0.012** (0.006)
Constant	2.304*** (0.533)	3.223*** (0.643)	7.796*** (0.577)	6.354*** (0.681)
Observations	283	283	283	283
Log Likelihood	−510.750	−509.655	−522.100	−519.575
Akaike Inf. Crit.	1,035.500	1,033.310	1,058.200	1,053.150
Bayesian Inf. Crit.	1,061.018	1,058.828	1,083.718	1,078.668
Note:				*p<0.1; **p<0.05; ***p<0.01

Self-reported scores were also regressed on the vehicle state predictors as with other response variables. There was only one significant main effect of the presence of a lead vehicle for both on-the-fly comfort ($b = -1.166, sd = 0.277, p < 0.05$) and anxiety ($b = 1.104, sd = 0.283, p < 0.05$) scores. There was also one significant interaction effect ($b = 1.359, sd = 0.631, p < 0.05$) of lead vehicle presence and the maximum longitudinal jerk on the comfort score but not on the anxiety scores.

There were no main or interaction effects for any vehicle state predictor on the CSAI-2 scores.

5.8 Qualitative Interview Feedback

As mentioned in Chapter 3, each experiment concluded with a brief interview in which participants were asked for feedback on the events they encountered, and their overall

Table 5.15: Effects of physiological response on CSAI-2 scores.

	Dependent variable:	
	CSAI2	
	(1)	(2)
NumPks	0.319*** (0.086)	0.304*** (0.085)
MaxPkAmp	−0.583 (0.515)	
MaxHr	−0.011 (0.011)	
MeanScl		−5.672 (8.197)
MeanHr		−0.016 (0.024)
Constant	11.732*** (1.905)	11.572*** (2.624)
Observations	80	80
Log Likelihood	−180.674	−181.283
Akaike Inf. Crit.	375.348	376.567
Bayesian Inf. Crit.	392.022	393.241
Note:		*p<0.1; **p<0.05; ***p<0.01

impression of the study procedure. Participants rated each of the four events in terms of how uncomfortable or anxious they felt.

5.8.1 Event-specific feedback

For most participants, stopping events were the most uncomfortable or anxiety-inducing (16 out of 20). Of the 16 participants, 14 rated stopping behind a car as the most uncomfortable, while 3 considered stopping at an intersection to be the most uncomfortable (1 participant considered both stops to be equally uncomfortable).

Likewise, for most participants, the turning event and passing another car did not have much of an impact. 11 of 20 participants were not affected by turns, while for 12 participants, passing the other car made no difference. In all, 7 out of 20 participants considered both turns and passing the other car to have little to no effect on their anxiety or discomfort.

There were a few participants, however, who were either quite or somewhat affected by turning and passing events. 8 out of 20 participants remarked that turning made them uncomfortable, while 4 participants expressed discomfort at passing the other car.

5.8.2 Additional comments

None of the participants experienced any motion sickness except for one during the last trial.

"Jerkiness" in motion

Some participants mentioned the motion of the car as a cause for their anxiety. Participant P10 mentioned that "jerkiness during breaking" and "revving" made them feel anxious, and said that "the drive would have been more comfortable if there was less jerkiness". Participant P17 described the driving as "a bit jittery" with an "abrupt but OK" stop at the intersection while P23 felt that the motion was too "jerky".

Meanwhile, P22 considered the driving to be outright "dangerous" and "bad", in reference to the car-stop event, while P24 likened the car motion to "when my kids were learning to drive".

Turns

The few participants who said that they were affected by turns were uncomfortable either by the sharpness or the trajectory of the turn. Even some of the participants who were not too affected by the turning described certain "oddities". For example, participant P16 was quoted as saying "The turns didn't affect me much, but I didn't really like them" while P17 said that the turns were "unnatural but didn't make me anxious". P17 further described their thoughts on the turn saying that it "looked like we were going off-road".

P22, who considered one of the turns to be "sharp but still relatively comfortable", although another trial left them feeling anxious when they felt that the car was possibly moving along the wrong trajectory. P23 was uncomfortable at the speed at which the turns were taken as well as by how sharp they were and felt that they were "hard to predict".

Participant P26 was worried about the possibility of the car overshooting the turn, while P27 was concerned that the car was "too close to the side".

Passing a parked car

While passing the parked car did not cause much discomfort except in only a few of the participants, some did express a level of attentiveness about the event: in particular, P13 said that the event made them more "curious" rather than uncomfortable or anxious about what the interaction with the parked car was going to be like. P17, on the other hand, felt that the "car was too close". Similar thoughts were expressed by P22 who said that "as a cyclist, I was worried when we came too close [to the parked car]", while P26 said that although they were uncomfortable at first, they "eventually got used to it".

5.8.3 Unexpected event occurrences

Despite the efforts taken to ensure consistency during the course of the experiment, there were at times unexpected occurrences that were beyond the control of the experiment. For example, there were a few times during which a manual takeover was necessary due to failure of the ADS - this was usually necessary during the car-stop event when the ADS failed to properly detect the lead vehicle in time. At other times, there were unexpected vehicles parked in the opposite lane despite the efforts to ensure that the track was closed during the course of the experiment. While these parked vehicles did not affect the route or planned sequence of the study, their presence had an effect on some participants.

Participants P11 and P12 experienced a spike in their anxiety level when they noticed the takeover that occurred during their trial, while P23 saw the takeover but wasn't affected by it at all. Meanwhile, P22 did not even notice the takeover.

During one of the trials for P14, there was an unexpected beeping alert that occurred when the ego vehicle approached the lead vehicle. Interestingly, this did not trigger any apprehension in the participant but instead was a source of reassurance (even though the alert was unexpected and did not really indicate any safety procedure to be undertaken).

Participant P28 was particularly affected by the parked motorcycles that were present during their trials. This was also reflected in their CSAI-2 score which was the highest for their first trial, but which lowered as they grew used to the presence of the vehicles.

Chapter 6

Discussion

A major aspect of this study is the fact that it was run in a real-world autonomous vehicle as opposed to the simulated environments prevalent in state-of-the-art research. This allowed for making informed conclusions on the actual impact of the driving style parameters considered, and also led to the generation of a naturalistic dataset of passenger response to driving style parameters.

We also contributed a method that could allow for greater control of complex, interdependent variables in future studies: through the use of independent events, we were able to isolate the different driving style parameters as much as realistically possible, even though these parameters may never be completely isolated due to the very nature of their combined role in realistic driving.

This rest of this chapter highlights all of the major experimental findings, relates them back to the qualitative feedback from participants, and proposes additional methods to improve the quality of the analysis as well as possible extensions to the work.

6.1 Experimental Findings

Our experimental findings allowed us to be able to address the research questions highlighted in Chapter 3:

1. *How does driving style correlate with passenger physiological response?*
2. *How does passenger physiological response correlate with self-reported comfort and anxiety?*

3. *How does driving style correlate with passenger self-reported comfort and anxiety?*

6.1.1 The role of events

One of the major findings of this experiment was the role of events. While the events were initially designed to isolate and test specific components of the variables and, by extension, type of driving style we were testing, it turned out that the type of event itself played a major role in determining what the individual response was going to be. Stopping events had the largest impact on response - both qualitatively as well as quantitatively, while turns and passing a parked car did not seem to elicit much of a response in general. Within the stopping events, the presence of a lead vehicle resulted in the highest overall response levels.

6.1.2 The role of driving style

This section answers our first research question, i.e., *"How does driving style correlate with passenger physiological response?"*

Acceleration and jerk

As expected, longitudinal acceleration and jerk had significant effects on most response variables, specifically, all the GSR variables and the maximum HR. As these variables are associated with decelerating to a stop and starting from a stop, the result was not surprising.

What was more interesting was how the direction of acceleration and jerk influenced the level of response. For the longitudinal components of acceleration and jerk, the positive direction was more influential than the negative direction (for example, acceleration had b values of 3.6 and 0.25 in the positive direction versus -1.9 and -0.08 in the negative direction, for the maximum HR and number of peaks). This suggests that people are more likely to be affected by positive forward acceleration rather than deceleration.

Lateral acceleration, on the other hand, showed the opposite effect: the negative direction was far more influential than the positive direction. This result was surprising due to the absence of left turns (the negative direction) during driving, and may be a result of noisy accelerometer readings or the presence of a confounding variable that we did not consider.

The moderating effect of longitudinal distance (to a lead vehicle)

We also saw that the mere presence of a lead vehicle had both a significant overall impact on response, as well as a moderating effect on the longitudinal acceleration and jerk components. By extension, in the presence of a lead vehicle, the longitudinal distance to the vehicle itself also had a similar moderating effect. More specifically, the closer the ego vehicle came to the lead vehicle, the higher was the effect of acceleration and jerk.

This result further bolsters our findings from Section 6.1.1, i.e., that the magnitude of response is context dependent and, in particular, is greatly influenced not just by the type of event but the mere presence of another interacting agent vehicle within the event.

6.1.3 Relating driving style and physiological response to comfort and anxiety

This section answers our second and third research questions, i.e., *"How does passenger physiological response correlate with self-reported comfort and anxiety?"* and *"How does driving style correlate with self-reported comfort and anxiety?"*

Physiological response and self-reported scores

A major goal of the study was to use the measured physiological response as an indicator of discomfort or anxiety in the passenger. While the chosen response variables are known, from literature, to be indicators of anxiety and discomfort, we wanted to further solidify this connection in our work and answer our second research question.

The analysis on both the standardized CSAI-2 questionnaire as well as the on-the-fly self-reported anxiety and comfort scores suggests that the number of peaks and the mean HR may indeed indicate the presence of anxiety or discomfort. For instance, for every unit increase in the number of peaks, the predicted comfort decreases by about a quarter of a point on the 10-point Likert scale, while the CSAI2 score increases by about 30%.

Driving style and self-reported scores

With respect to the third research question, there was a significant effect of the presence of a lead vehicle on the self-reported comfort and anxiety scores. Specifically, the presence of a lead vehicle leads to a 1.2 times decrease in the on-the-fly comfort score and a 1.1

times increase in the on-the-fly anxiety score. An interaction effect between this presence and longitudinal jerk also existed in which the effect of positive jerk, for instance, was magnified nearly 1.4 times by the presence of the lead vehicle. These results suggests a possible correlation of driving style with the on-the-fly scores.

Interestingly, there was no significant effect on the CSAI-2 scores. While this is unexpected (due to the presence of an effect for the physiological responses), it is probably due to the fact that different vehicle state variables play a larger role at different events resulting in non-representative aggregations at the trial level; the physiological response, on the other hand, is more consistent even when aggregated over a trial.

6.2 The Bigger Picture: Relating Quantitative Results with Qualitative Findings

Most of the results from the statistical analysis are directly mirrored in the interview responses from participants:

- **Stopping events and longitudinal distance**
  Qualitative: The majority of participants found stopping events to be the most uncomfortable or anxiety-inducing. Moreover, most of these participants considered stopping behind a car to be more uncomfortable than stopping at a clear intersection.

  Quantitative: Stopping events had the most significant physiological response, in general, and car-stops had a greater effect than intersection stops. In addition, the presence of a lead vehicle as well as its distance from the ego vehicle significantly moderated the effect of acceleration and jerk.

- **Jerkiness in motion**
  Qualitative: Some participants were uncomfortable with the "jerkiness" in motion while driving, particularly when decelerating to and accelerating from a stop.

  Quantitative: There was an overall positive effect of the magnitude of longitudinal jerk on the response variables. In addition, this effect was greater for the positive direction.

- **Turning**
  Qualitative: For the most part, participants were generally not affected by turns, although some did consider them to be sharp.

Quantitative: Lateral acceleration had a significant effect only with respect to heart rate, but not for any other response variable.

- **Passing a car**
  Qualitative: Only a fifth of the participants felt discomfort when passing the parked car.

  Quantitative: The presence of a parked car did not have a significant effect for any response variable.

6.3 Applications in the Real World

As our findings indicate a possible association between the chosen response variables and comfort and anxiety scores, the predictor variables that were found to be significant could potentially be used in an implicit interface between the passenger and the vehicle. The goal of the ADS would then be to implicitly sense the onset of anxiety or discomfort both through the level of each physiological response considered, as well as from the context represented by the vehicle state variables (the significant predictors). Based on this sensed affective state, the ADS could then automatically adapt the driving style to one that induces less discomfort or anxiety in the passenger.

Of course, in the real world, it would be highly impractical to set up passengers with multiple sensors each time they ride in an autonomous vehicle. It would then be interesting to see how the study results can be adapted for a system that employs a set of unobtrusive sensing techniques: perhaps a stationary eye-tracker could be fixed in front of the passenger in order to measure eye movement patterns, while heart rate and GSR could be communicated by means of wearable devices such as smartwatches or rings.

On the other hand, instead of a more complex and general adaptive system, perhaps future autonomous vehicles could simply incorporate driving profiles at an individual level, wherein each passenger would have a profile tailored to their observed range of responses for different driving contexts and styles.

For both of these proposed solutions, an immediate concern that comes up, however, is the multiple passenger problem - how could the driving style be tailored to multiple passengers with different driving preferences? One solution could be to use existing data to estimate the most general zone of comfort for each vehicle state variable considered and use this as a default setting when an individual passenger's needs cannot be feasibly met. Another might involve optimizing on-the-fly, the general profile by considering only the profiles of the passengers present in the vehicle at the same time.

6.4 Future Work

There is much scope to expand and improve the current experimental design and analysis. This section outlines a few possible ways to do this.

6.4.1 Comparison with a human driver

This study did not compare the responses to those that would have been obtained with a human driver. It would be beneficial to run an extension to this work by conducting a between-subjects experiment in which one group was made to experience the same driving style but carried out by a human driver as opposed to the ADS. To perfectly imitate the driving style of the ADS, the experiment could be run with deception in which the human driver merely has their hands on the steering wheel but does not actually drive.

6.4.2 Data

It would be useful to expand the dataset to include samples from a larger pool of participants; 20 participants might not be sufficient to truly capture variation in driving preferences and responses. We have already seen that while turns and passing did not affect the majority of participants, these events still had a substantial impact on a few individuals and was even regarded by them as the most uncomfortable of all four events.

6.4.3 Exploring additional variables

The only variables considered as predictors in this study were vehicle state variables like acceleration and distance. However, it would probably be far more beneficial to the analysis if control variables were added which would help explain a greater portion of the individual differences in response. Such control variables could include the age and sex of participants, as well as individual biases such as preferred driving style, and familiarity with autonomous vehicles.

Besides controlling for age, gender, and personal biases, it might also be more helpful to consider different vehicle state variables themselves. Perhaps the variables chosen in this study may not have been as significant in prediction as the other potential choices. For instance, the longitudinal distance could be replaced with time to collision, and velocity of the ego vehicle could be an additional predictor.

Weather could also be an additional control variable: perhaps participants feel more at ease regarding their safety in the autonomous vehicle when it is bright and sunny (optimal conditions for the ADS to function) as opposed to when it is raining. There was considerable variation in weather over the week in which the experiment was run, making this variable a suitable candidate for inclusion in future analyses.

6.4.4 Accounting for unexpected events

There were a few unexpected events that were not taken into account for the statistical analysis - for instance, the presence of parked motorcycles on a closed test track, and the occasional takeovers that occurred when the ADS system was unable to detect the lead vehicle. Such events may have had a significant impact on the response of the individuals involved. It would be interesting to conduct a separate analysis on these unexpected events to see if the responses were indeed affected.

6.4.5 Improving the models used

The analysis conducted so far considered only an aggregation-based approach wherein each variable was aggregated over an interval that spanned either the entire trial, or 5 second non-overlapping windows. Due to this, there is necessarily some amount of data loss as a result of the aggregation.

In order to overcome this issue, a detailed time-series approach could be used in which each sample is considered on its own instead of aggregating samples over a specified interval. Such an analysis, while being more involved and complex, would be likely to give us a better, more accurate, picture of the effect of each driving style parameter.

Chapter 7

Conclusion

In this study we ran an experiment in a physical autonomous vehicle that required passengers to experience a sequence of driving styles ranging from defensive to aggressive. We collected physiological measurements (skin response, heart rate variation, and eye movement patterns) along with self-reported comfort and anxiety scores to gain a holistic perspective on passenger response to driving style in autonomous vehicles. The fact that this experiment was run in a real-world autonomous vehicle and not in a simulated environment, is one of the major contributions of this work.

We analyzed driving style parameters at two levels of aggregation: over an entire trial, and over 5 second consecutive, non-overlapping windows. We also conducted separate analyses on the role that events played – irrespective of driving style parameters – on passenger response, and on the association of physiological response and driving style with self-reported scores.

Longitudinal distance, acceleration, and jerk, were found to have significant effects on the response variables considered, with positive acceleration and jerk having larger effects than their negative counterparts. Moreover, the magnitude of a response is situation dependent: both the presence of a lead vehicle and its proximity to the ego significantly moderate the effect of acceleration and jerk. Events also played a role in affecting the overall response. Stopping events generated a higher level of response than non-stopping events, with stopping behind a car having the largest effect. These findings were consistent with qualitative feedback from the participants, most of whom felt the greatest sense of discomfort and anxiety when stopping behind the lead vehicle.

Finally, the analysis between physiological response and the self-reported scores shows that the responses could potentially be used as an indirect indicator of anxiety or discomfort, a finding consistent with the surveyed literature.

References

[1] Pieter Abbeel and Andrew Y Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*, page 1. ACM, 2004.

[2] Irman Abdic, Lex Fridman, Daniel McDuff, Erik Marchi, Bryan Reimer, and Björn Schuller. Driver frustration detection from audio and video in the wild. In *KI 2016: Advances in Artificial Intelligence: 39th Annual German Conference on AI, Klagenfurt, Austria, September 26-30, 2016, Proceedings*, volume 9904, page 237. Springer, 2016.

[3] Jonathan Allsop, Rob Gray, Heinrich H Bülthoff, and Lewis Chuang. Eye movement planning on single-sensor-single-indicator displays is vulnerable to user anxiety and cognitive load. *Journal of Eye Movement Research*, 10(5):8–1, 2017.

[4] Georg W Alpers, Frank H Wilhelm, and Walton T Roth. Psychophysiological assessment during exposure in driving phobic patients. *Journal of Abnormal Psychology*, 114(1):126, 2005.

[5] Monica Babes, Vukosi Marivate, Kaushik Subramanian, and Michael L Littman. Apprenticeship learning about multiple intentions. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 897–904, 2011.

[6] Chandrayee Basu, Qian Yang, David Hungerman, Mukesh Singhal, and Anca D. Dragan. Do you want your autonomous car to drive like you? In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '17, pages 417–425, New York, NY, USA, 2017. ACM.

[7] Philipp Bender, Julius Ziegler, and Christoph Stiller. Lanelets: Efficient map representation for autonomous driving. In *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pages 420–425. IEEE, 2014.

[8] Gerard J Blaauw. Driving experience and task demands in simulator and instrumented car: a validation study. *Human Factors*, 24(4):473–486, 1982.

[9] Jens Blechert, Marta Lajtman, Tanja Michael, Jürgen Margraf, and Frank H Wilhelm. Identifying anxiety states using broad sampling and advanced processing of peripheral physiological information. *Biomedical sciences instrumentation*, 42:136–141, 2006.

[10] Pieter Blignaut. Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, 71(4):881–895, 2009.

[11] Shadan Sadeghian Borojeni, Torben Wallbaum, Wilko Heuten, and Susanne Boll. Comparing shape-changing and vibro-tactile steering wheels for take-over requests in highly automated driving. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '17, pages 221–225, New York, NY, USA, 2017. ACM.

[12] Joe Causer, Paul S Holmes, Nickolas C Smith, and A Mark Williams. Anxiety, movement kinematics, and visual attention in elite-level performers. *Emotion*, 11(3):595, 2011.

[13] Calais KY Chan and Peter F Lovibond. Expectancy bias in trait anxiety. *Journal of abnormal psychology*, 105(4):637, 1996.

[14] Jeffrey Cohn, Takeo Kanade, Tsuyoshi Moriyama, Zara Ambadar, Jing Xiao, Jiang Gao, and Hiroki Imamura. A comparative study of alternative facs coding algorithms. 05 2002.

[15] SAE On-Road Automated Driving Committee et al. Sae j3016. taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. Technical report, tech. rep., SAE International, 2016. Cited on.

[16] Joost CF De Winter, Riender Happee, Marieke H Martens, and Neville A Stanton. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation research part F: traffic psychology and behaviour*, 27:196–217, 2014.

[17] Forward Development. City car driving. https://citycardriving.com/, 2001. Accessed: 2019-09-11.

[18] Thomas A Dingus, Feng Guo, Suzie Lee, Jonathan F Antin, Miguel Perez, Mindy Buchanan-King, and Jonathan Hankey. Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, 113(10):2636–2641, 2016.

[19] Will autonomous vehicles be accessible to people with disabilities?, 2017. https://sharedusemobilitycenter.org/news/will-autonomous-vehicles-be-accessible-to-people-with-disabilities/.

[20] Weishan Dong, Jian Li, Renjie Yao, Changsheng Li, Ting Yuan, and Lanjun Wang. Characterizing driving styles with deep learning. *arXiv preprint arXiv:1607.03611*, 2016.

[21] Sidney K Dmello and Arthur Graesser. Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction*, 20(2):147–187, 2010.

[22] Stephen Edelstein. Kia wants future autonomous cars to be able to read passengers emotions. https://www.digitaltrends.com/cars/kia-emotion-recognition-ai-tech-at-ces-2019/, 2018. Accessed: 2019-09-15.

[23] Ellen Edmonds. Three in four americans remain afraid of fully self-driving vehicles.

[24] Paul Ekman. Facial action coding system. 1977.

[25] Michael Festner, Alexandra Eicher, and D Schramm. Beeinflussung der komfort- und sicherheitswahrnehmung beim hochautomatisierten fahren durch fahrfremde tätigkeiten und spurwechseldynamik. *von*, 11, 2017.

[26] Lex Fridman, Heishiro Toyoda, Sean Seaman, Bobbie Seppelt, Linda Angell, Joonbum Lee, Bruce Mehler, and Bryan Reimer. What can be predicted from six seconds of driver glances? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2805–2813. ACM, 2017.

[27] David Goedicke, Jamy Li, Vanessa Evers, and Wendy Ju. Vr-oom: Virtual reality on-road driving simulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 165. ACM, 2018.

[28] Peter Grant, Bruce Artz, Jeff Greenberg, and Larry Cathey. Motion characteristics of the virttex motion system. In *Proceedings of the 1st human-centered transportation simulation conference*, pages 4–7, 2001.

[29] Mohammed Hoque and Rosalind W Picard. Acted vs. natural frustration and delight: Many people smile in natural frustration. In *Face and Gesture 2011*, pages 354–359. IEEE, 2011.

[30] Lynn M Hulse, Hui Xie, and Edwin R Galea. Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety science*, 102:1–13, 2018.

[31] iMotions. Galvanic skin response (gsr): The complete pocket guide. https://imotions.com/blog/galvanic-skin-response/, 2016. Accessed: 2019-09-11.

[32] Myounghoon Jeon. Multimodal displays for take-over in level 3 automated vehicles while playing a game. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, page LBW0229. ACM, 2019.

[33] Takeo Kanade, Jeffrey F Cohn, and Yingli Tian. Comprehensive database for facial expression analysis. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pages 46–53. IEEE, 2000.

[34] Vassilios Krassanakis, Vassiliki Filippakopoulou, and Byron Nakos. Eyemmv toolbox: An eye movement post-analysis tool based on a two-step spatial dispersion threshold for fixation identification. 2014.

[35] Jason Ku, Melissa Mozifian, Jungwook Lee, Ali Harakeh, and Steven L Waslander. Joint 3d proposal generation and object detection from view aggregation. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–8. IEEE, 2018.

[36] Markus Kuderer, Shilpa Gulati, and Wolfram Burgard. Learning driving styles for autonomous vehicles from demonstration. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 2641–2646. IEEE, 2015.

[37] Michael Glazer William Angell Spencer Dodd Benedikt Jenik Jack Terwilliger Aleksandr Patsekin Julia Kindelsberger Li Ding Sean Seaman Alea Mehler Andrew Sipperley Anthony Pettinato Bobbie Seppelt Linda Angell Bruce Mehler Bryan Reimer Lex Fridman, Daniel E. Brown. Mit advanced vehicle technology study: Large-scale naturalistic driving study of driver behavior and interaction with automation. *IEEE Access*, 7:102021–102038, 2019.

[38] Changchun Liu, Pramila Agrawal, Nilanjan Sarkar, and Shuo Chen. Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *International Journal of Human-Computer Interaction*, 25(6):506–529, 2009.

[39] Rainer Martens, Damon Burton, Robin S Vealey, Linda A Bump, and Daniel E Smith. Development and validation of the competitive state anxiety inventory-2. *Competitive anxiety in sport*, pages 117–190, 1990.

[40] J Morales, V Garcia, X García-Massó, P Salvá, R Escobar, et al. The use of heart rate variability in assessing precompetitive stress in high-standard judo athletes. *International journal of sports medicine*, 34(02):144–151, 2013.

[41] Nadia Mullen, Judith Charlton, Anna Devlin, and Michel Bedard. *Simulator validity: behaviours observed on the simulator and on the road*, pages 1 – 18. CRC Press, Australia, 1st edition, 2011.

[42] Hiroki Murakami and Hideki Ohira. Influence of attention manipulation on emotion and autonomic responses. *Perceptual and Motor Skills*, 105(1):299–308, 2007.

[43] Sonia Baltodano Nikolas Martelaro Rohan Maheshwari David Miller Wendy Ju Nikhil Gowda, Srinath Sibi. Nudge: Haptic pre-cueing to communicate automotive intent. In *Automotive User Interfaces (AutomotiveUI), 2015 ACM*. ACM, 2015.

[44] Tomi Nukarinen, Jussi Rantala, Ahmed Farooq, and Roope Raisamo. Delivering directional haptic cues through eyeglasses and a seat. In *World Haptics Conference (WHC), 2015 IEEE*, pages 345–350. IEEE, 2015.

[45] Simone Pettigrew, Caitlin Worrall, Zenobia Talati, Lin Fritschi, and Richard Norman. Dimensions of attitudes to autonomous vehicles. *Urban, Planning and Transport Research*, 7(1):19–33, 2019.

[46] Rosalind W Picard. Affective computing-mit media laboratory perceptual computing section technical report no. 321. *Cambridge, MA*, 2139, 1995.

[47] Imotions: Biometric Research Platform. Eye tracking: The complete pocket guide.

[48] Christopher J Ploch, Jung Hwa Bae, Caitlin C Ploch, Wendy Ju, and Mark R Cutkosky. Comparing haptic and audio navigation cues on the road for distracted drivers with a skin stretch steering wheel. In *World Haptics Conference (WHC), 2017 IEEE*, pages 448–453. IEEE, 2017.

[49] Kathrin Pollmann, Oilver Stefani, Amelie Bengsch, Matthias Peissner, and Mathias Vukelić. How to work in the car of the future?: A neuroergonomical study assessing concentration, performance and workload based on subjective, behavioral and neurophysiological insights. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 54:1–54:14, New York, NY, USA, 2019. ACM.

[50] Pramila Rani, Nilanjan Sarkar, and Julie Adams. Anxiety-based affective communication for implicit human–machine interaction. *Advanced Engineering Informatics*, 21(3):323–334, 2007.

[51] Pramila Rani, Nilanjan Sarkar, Craig A Smith, and Leslie D Kirby. Anxiety detecting robotic system–towards implicit human-robot collaboration. *Robotica*, 22(1):85–95, 2004.

[52] Relaxnews. The toyota concept-i gives ai a human face. https://www.ctvnews.ca/autos/the-toyota-concept-i-gives-ai-a-human-face-1.3230369, 2017. Accessed: 2019-09-15.

[53] Thomas Ritz, Andrew Steptoe, Stephen DeWilde, and Marco Costa. Emotions and stress increase respiratory resistance in asthma. *Psychosomatic Medicine*, 62(3):401–412, 2000.

[54] Felix Ros, Jacques Terken, Frank van Valkenhoef, Zane Amiralis, and Stefan Beckmann. Scribble your way through traffic. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '18, pages 230–234, New York, NY, USA, 2018. ACM.

[55] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In *2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN)*, pages 795–802. IEEE, 2016.

[56] Henry A Rowley. Neural network-based face detection. Technical report, CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE, 1999.

[57] Aitor Ruano. Deepgtav. https://github.com/aitorzip/DeepGTAV, 2017. Accessed: 2019-09-11.

[58] Shadan Sadeghian Borojeni, Susanne CJ Boll, Wilko Heuten, Heinrich H Bülthoff, and Lewis Chuang. Feel the movement: Real motion influences responses to take-over requests in highly automated vehicles. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 246. ACM, 2018.

[59] Dario D. Salvucci and Joseph H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, ETRA '00, pages 71–78, New York, NY, USA, 2000. ACM.

[60] Albrecht Schmidt. Implicit human computer interaction through context. *Personal technologies*, 4(2-3):191–199, 2000.

[61] Albrecht Schmidt, H-W Gellersen, and Christian Merz. Enabling implicit human computer interaction: a wearable rfid-tag reader. In *Digest of Papers. Fourth International Symposium on Wearable Computers*, pages 193–194. IEEE, 2000.

[62] Thibaud Senechal, Daniel McDuff, and Rana Kaliouby. Facial action unit detection using active learning and an efficient non-linear kernel approximation. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 10–18, 2015.

[63] Michael Sivak and Brandon Schoettle. Universal motion sickness countermeasure system, Apr 2017.

[64] Charles D Spielberger. Theory and research on anxiety. pages 3–19. New York: Academic Press, 1966.

[65] David L Strayer, Joel M Cooper, Jonna Turrill, James R Coleman, and Rachel J Hopman. Measuring cognitive distraction in the automobile iii: A comparison of ten 2015 in-vehicle information systems. 2015.

[66] Hai Tao and Thomas S Huang. Explanation-based facial motion tracking using a piecewise bezier volume deformation model. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 1, pages 611–617. IEEE, 1999.

[67] Orit Taubman-Ben-Ari, Mario Mikulincer, and Omri Gillath. The multidimensional driving style inventoryscale construct and validation. *Accident Analysis & Prevention*, 36(3):323–332, 2004.

[68] Ariel Telpaz, Brian Rhindress, Ido Zelman, and Omer Tsimhoni. Haptic seat for automated driving: Preparing the driver to take control effectively. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '15, pages 23–30, New York, NY, USA, 2015. ACM.

[69] Jacques Terken, Pierre Levy, Chao Wang, Juffrizal Karjanto, Nidzamuddin Md Yusof, Felix Ros, and Sergej Zwaan. Gesture-based and haptic interfaces for connected and autonomous driving. In *Advances in Human Factors and System Interactions*, pages 107–115. Springer, 2017.

[70] Sebastian Thrun, Mike Montemerlo, Hendrik Dahlkamp, David Stavens, Andrei Aron, James Diebel, Philip Fong, John Gale, Morgan Halpenny, Gabriel Hoffmann, et al. Stanley: The robot that won the darpa grand challenge. *Journal of field Robotics*, 23(9):661–692, 2006.

[71] Anna Tobin. Jaguar land rover makes breakthrough in cure for motion sickness. https://www.forbes.com/sites/annatobin/2018/11/02/jaguar-land-rover-makes-breakthrough-in-cure-for\-motion-sickness/#105358a565c0, 2018. Accessed: 2019-09-15.

[72] Toyota. Driving simulator. http://www.toyota.com.cn/innovation/safety_technology/safety_measurements/driving_simulator.html, 2007. Accessed: 2019-09-11.

[73] Remo van der Heiden, Shamsi T Iqbal, and Christian P Janssen. Priming drivers before handover in semi-autonomous cars. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 392–404. ACM, 2017.

[74] Himanshu Verma, Florian Evéquoz, Guillaume Pythoud, Grace Eden, and Denis Lalanne. Engaging pedestrians in designing interactions with autonomous vehicles. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, pages LBW0126:1–LBW0126:6, New York, NY, USA, 2019. ACM.

[75] Paul Viola and Michael J Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.

[76] Marcel Walch, Tobias Sieber, Philipp Hock, Martin Baumann, and Michael Weber. Towards cooperative driving: Involving the driver in an autonomous vehicle's decision making. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 261–268. ACM, 2016.

[77] Ying Wang, Bruce Mehler, Bryan Reimer, Vincent Lammers, Lisa A D'Ambrosio, and Joseph F Coughlin. The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics*, 53(3):404–420, 2010.

[78] Mark Weiser. The computer for the 21 st century. *Scientific american*, 265(3):94–105, 1991.

[79] Mark R Wilson, Samuel J Vine, and Greg Wood. The influence of anxiety on visual attentional control in basketball free throw shooting. *Journal of Sport and Exercise Psychology*, 31(2):152–168, 2009.

[80] Mark R Wilson, Greg Wood, and Samuel J Vine. Anxiety, attentional control, and performance impairment in penalty kicks. *Journal of Sport and Exercise Psychology*, 31(6):761–775, 2009.

[81] Priscilla NY Wong, Duncan P Brumby, Harsha Vardhan Ramesh Babu, and Kota Kobayashi. " watch out!": Semi-autonomous vehicles using assertive voices to grab distracted drivers' attention. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, page LBW2312. ACM, 2019.

[82] Li Xu, Jie Hu, Hong Jiang, and Wuqiang Meng. Establishing style-oriented driver models by imitating human driving behaviors. *IEEE Transactions on Intelligent Transportation Systems*, 16(5):2522–2530, 2015.

[83] Thorsten O Zander, Jonas Brönstrup, Romy Lorenz, and Laurens R Krol. Towards bci-based implicit control in human–computer interaction. In *Advances in Physiological Computing*, pages 67–90. Springer, 2014.

[84] Zhihong Zeng, Jilin Tu, Brian Pianfetti, Ming Liu, Tong Zhang, Zhenqiu Zhang, Thomas S Huang, and Stephen Levinson. Audio-visual affect recognition through multi-stream fused hmm for hci. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 967–972. IEEE, 2005.

[85] Yali Zheng, Tracy CH Wong, Billy HK Leung, and Carmen CY Poon. Unobtrusive and multimodal wearable sensing to quantify anxiety. *IEEE Sensors Journal*, 16(10):3689–3696, 2016.

[86] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI*, volume 8, pages 1433–1438. Chicago, IL, USA, 2008.

APPENDICES

Appendix A

Study Materials

The next few pages contain the consent form, questionnaires, and remuneration form used during the study. There were some questions asked in the pre-study questionnaire that were not used in this work, particularly about experience and knowledge about autonomous vehicles, and preferred driving style - this data forms the basis for our future work.

Consent for Participation

By providing your consent, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

Title of the study: Passenger-vehicle interaction in Autonomous Vehicles

I have read the information presented in the information letter about a study conducted by Dr. Krzysztof Czarnecki, Dr. Oliver Schneider, Dr. Edith Law and Nicole Dillen at the Cheriton School of Computer Science, University of Waterloo, and the Department of Electrical and Computer Engineering, University of Waterloo. I have had the opportunity to ask questions related to the study and have received satisfactory answers to my questions and any additional details. I was informed that participation in the study is voluntary and that I can withdraw this consent by informing the researcher.

By signing below, I consent to my participation in this study designed to help to understand the effect of different driving styles employed by an autonomously driven vehicle on the anxiety levels of a passenger. I have read the letter of information and understand the risks and benefits of participation. I also understand that:

- My identity and data will be kept anonymous and confidential
- I agree to my study session being video recorded for the purpose of tracking my facial cues for analysis purposes
- I agree to my study session being audio recorded for the purpose of analyzing my audio responses during the course of the experiment
- My anonymized data may be used in future studies and possibly in a public open-access repository
- I agree to the use of anonymous quotations in any thesis or publication that comes from this research
- I am free to withdraw from the experiment at any time before or during, but not after, without reason or consequence
- I have been told the purpose of the experiment and am free to ask questions at any time
- I may take any questions regarding the results of the study and procedures followed to the investigators
- I may take any complaints or concerns to the Office of Research Ethics.

Please indicate below whether you agree to an anonymized or non-anonymized version of your video being shared publicly. The anonymized version would have your face blurred, and voice altered or muted. In the non-anonymized version, your face will be visible and your voice will be heard. In both cases, your name would not be associated with the video.

I consent to having an **anonymized** version of my video shared:
__ at an academic dissemination (such as a conference)
__ as a video on social media platforms (such as YouTube)

I consent to having a **non-anonymized** version of my video shared:
__ at an academic dissemination (such as a conference)
__ as a video on social media platforms (such as YouTube)

I have read the above statement and freely consent to participate in this research.

Participant's Name: ____

Participant's Signature: _____________________ Date: _________

Pre-Study Demographics Questionnaire

(1) How old are you? _________

(2) What is your gender? **Male / Female / Other / Prefer not to say**

(3a) If you are a student, what faculty are you in? What is your major?
Faculty: ________________________________ Major: ________________

(3b) If you are not a student, what is your area of work?

____

(4) The following questions relate to your experience as a **passenger** in an everyday driving scenario. Please circle the appropriate number - there are no right or wrong answers.

	Not at all	Somewhat	Moderately	Very much
a. I feel unsafe at high speeds.	1	2	3	4
b. I get uncomfortable with heavier congestion.	1	2	3	4
c. Getting too close to other vehicles makes me feel uncomfortable.	1	2	3	4

d. My preferred driving style** is: **Very Defensive/ Defensive/ Aggressive/ Very Aggressive**

**Driving style or behavior can either be characterized by higher speeds, sharper turns and minimal headway with other vehicles, i.e., aggressive behavior, or by very low speeds and large time and distance headways, i.e., defensive behavior.

(5) Have you taken or are currently taking a course that deals with (partially or completely) autonomous vehicles? **(Yes / No)**

(6) Have you ever ridden in an autonomous vehicle? **(Yes / No)**

(7) Do you follow articles, blogs, podcasts, or stories related to autonomous vehicles on popular media? **(Yes / No)**

(8) On a scale of 1-10 (1 being **None at All** and 10 being **Highly Knowledgeable**) please rate your overall knowledge of autonomous vehicles. / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 /

(9) On a scale of 1-10 (1 being **Not Trusting** and 10 being **Fully Trusting**) please rate your overall trust in autonomous vehicles and their safety. / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 /

CSAI-2

Directions: Please read each statement and circle the appropriate number to indicate how you felt during the drive. There are no right or wrong answers. Do not spend too much time on any one statement.

	Not at all	Somewhat	Moderately	Very much
1. I felt nervous.	1	2	3	4
2. I felt jittery.	1	2	3	4
3. My body felt tense.	1	2	3	4
4. I felt tense in my stomach.	1	2	3	4
5. My body felt relaxed.	1	2	3	4
6. My heart was racing.	1	2	3	4
7. I felt my stomach sinking.	1	2	3	4
8. My hands were clammy.	1	2	3	4
9. My body felt tight.	1	2	3	4

University of Waterloo
Research Participant's Acknowledgement of
Receipt of Remuneration
and
Self-Declared Income

Section A: To be completed by Principal Investigator or designate

Principal/Faculty Investigator's Name: Krzysztof Czarnecki, Oliver Schneider

Student Investigator(s)'s Name: Nicole Dillen

Department: Computer Science

Study Title: Passenger-vehicle Interaction in Autonomous Vehicles

---

Section B: To be completed by research participant

In appreciation of my involvement as a research participant in the above study,
I acknowledge that I have received $ __________ from the University of Waterloo.

I further acknowledge that:

- this amount received from the University of Waterloo is taxable;
- that it is my responsibility to report the amount received for income tax purposes; and
- the University of Waterloo will not issue a tax receipt for the amount received.

Participant's Name: ________________________________

Participant's Signature: ________________________________

Date: ________________________________

Witness' Name ________________________________

Witness' Signature: ________________________________

Date: ________________________________

Appendix B

Detailed Tables

The tables used in the main thesis content had standard errors for each beta term instead of confidence intervals. In addition, for brevity, we did not report the complete table of effects when studying interaction terms. In this section, the complete list of effects along with confidence intervals are reported in detailed tables.

Table B.1: Main and interaction effects for Maximum Peak Amplitude at the window level.

	Dependent variable:	
	Max. Pk. Amp.	
	b $(lower, upper)$	b $(lower, upper)$
t	−0.002 (−0.006, 0.002)	−0.003 (−0.007, 0.001)
$present_{lead}$	0.121*** (0.048, 0.193)	−0.451** (−0.796, −0.106)
a^{max}_{long}	−0.026 (−0.063, 0.011)	−0.034 (−0.078, 0.010)
$dir_{a^{max}_{long}}$	−0.077*** (−0.130, −0.024)	−0.068** (−0.123, −0.013)
a^{max}_{lat}	0.036 (−0.160, 0.232)	0.061 (−0.136, 0.258)
$dir_{a^{max}_{lat}}$	−0.029 (−0.102, 0.044)	−0.047 (−0.122, 0.027)
$\dot{a}^{max}_{long}$	−0.025 (−0.067, 0.018)	−0.028 (−0.082, 0.025)
$dir_{\dot{a}^{max}_{long}}$	−0.013 (−0.044, 0.017)	−0.011 (−0.043, 0.021)
$\dot{a}^{max}_{lat}$	−0.024 (−0.095, 0.046)	−0.014 (−0.085, 0.057)
$\mathrm{dir}_{\dot{a}^{max}_{lon}}$	0.039** (0.003, 0.076)	0.038** (0.002, 0.074)
$present_{pass}$	−0.033 (−0.091, 0.024)	−0.036 (−0.093, 0.021)
a^{max}_{long}:$dir_{a^{max}_{long}}$	0.114*** (0.047, 0.182)	0.107*** (0.030, 0.184)
a^{max}_{lat}:$dir_{a^{max}_{lat}}$	−0.024 (−0.223, 0.175)	−0.048 (−0.248, 0.151)
$\dot{a}^{max}_{long}$:$dir_{\dot{a}^{max}_{long}}$	0.111*** (0.036, 0.186)	0.093** (0.003, 0.183)
$\dot{a}^{max}_{lat}$:$dir_{\dot{a}^{max}_{lat}}$	−0.033 (−0.157, 0.091)	−0.051 (−0.176, 0.074)
$present_{lead}1$:a^{max}_{long}		−0.147** (−0.263, −0.030)
$present_{lead}1$:$dir_{a^{max}_{long}}$		0.337 (−0.120, 0.793)
$present_{lead}1$:$\dot{a}^{max}_{long}$		−0.048 (−0.144, 0.047)
$present_{lead}1$:$dir_{\dot{a}^{max}_{long}}$		−0.001 (−0.229, 0.226)
$present_{lead}1$:a^{max}_{long}:$dir_{a^{max}_{long}}$		0.249** (0.059, 0.438)
$present_{lead}1$:$\dot{a}^{max}_{long}$:$dir_{\dot{a}^{max}_{long}}$		0.122* (−0.017, 0.261)
Constant	0.009 (−0.087, 0.106)	0.032 (−0.066, 0.129)
Observations	1,700	1,700
Log Likelihood	−208.368	−198.622
Akaike Inf. Crit.	456.736	449.244
Bayesian Inf. Crit.	565.504	590.642
Note:		*p<0.1; **p<0.05; ***p<0.01

Table B.2: Main and interaction effects for Number of Peaks at the window level.

	Dependent variable:	
	Num. Pks.	
	b $(lower, upper)$	b $(lower, upper)$
t	−0.001 (−0.008, 0.006)	−0.004 (−0.011, 0.003)
$present_{lead}$	0.432*** (0.271, 0.593)	−1.372*** (−2.159, −0.585)
a_{long}^{max}	−0.084* (−0.169, 0.001)	−0.137*** (−0.238, −0.035)
$dir_{a_{long}^{max}}$	−0.094* (−0.202, 0.014)	−0.013 (−0.127, 0.101)
a_{lat}^{max}	−0.002 (−0.436, 0.432)	0.101 (−0.331, 0.534)
$dir_{a_{lat}^{max}}$	−0.028 (−0.188, 0.133)	−0.072 (−0.233, 0.089)
$\dot{a}_{long}^{max}$	−0.066 (−0.161, 0.029)	0.009 (−0.111, 0.129)
$dir_{\dot{a}_{long}^{max}}$	−0.022 (−0.098, 0.054)	−0.047 (−0.127, 0.032)
$\dot{a}_{lat}^{max}$	−0.104 (−0.274, 0.066)	−0.090 (−0.260, 0.080)
$dir_{\dot{a}_{lat}^{max}}$	0.029 (−0.060, 0.118)	0.021 (−0.066, 0.109)
$present_{pass}$	−0.041 (−0.156, 0.075)	−0.043 (−0.157, 0.071)
$a_{long}^{max}{:}dir_{a_{long}^{max}}$	0.330*** (0.180, 0.479)	0.277*** (0.107, 0.448)
$a_{lat}^{max}{:}dir_{a_{lat}^{max}}$	0.018 (−0.422, 0.459)	−0.082 (−0.519, 0.355)
$\dot{a}_{long}^{max}{:}dir_{\dot{a}_{long}^{max}}$	0.333*** (0.167, 0.499)	0.199* (−0.005, 0.404)
$\dot{a}_{lat}^{max}{:}dir_{\dot{a}_{lat}^{max}}$	0.056 (−0.232, 0.344)	0.058 (−0.230, 0.347)
$present_{lead}1{:}a_{long}^{max}$		−0.381*** (−0.642, −0.119)
$present_{lead}1{:}dir_{a_{long}^{max}}$		0.137 (−0.905, 1.180)
$present_{lead}1{:}\dot{a}_{long}^{max}$		−0.334*** (−0.555, −0.114)
$present_{lead}1{:}dir_{\dot{a}_{long}^{max}}$		0.270 (−0.255, 0.796)
$present_{lead}1{:}a_{long}^{max}{:}dir_{a_{long}^{max}}$		1.134*** (0.715, 1.553)
$present_{lead}1{:}\dot{a}_{long}^{max}{:}dir_{\dot{a}_{long}^{max}}$		0.504*** (0.193, 0.815)
Constant	0.183 (−0.042, 0.408)	0.258** (0.032, 0.484)
Observations	1,700	1,700
Log Likelihood	−1,580.307	−1,549.804
Akaike Inf. Crit.	3,200.614	3,151.608
Bayesian Inf. Crit.	3,309.382	3,293.006
Note:		*p<0.1; **p<0.05; ***p<0.01

Table B.3: Main and interaction effects for Mean SCL at the window level.

	Dependent variable:	
	Mean SCL	
	b $(lower, upper)$	b $(lower, upper)$
t	0.00005 (−0.003, 0.003)	0.001 (−0.002, 0.004)
$present_{lead}$	0.067*** (0.025, 0.109)	−0.006 (−0.140, 0.127)
$\bar{a}_{long}$	0.023 (−0.029, 0.075)	−0.011 (−0.068, 0.045)
$\bar{a}_{lat}$	−0.002 (−0.035, 0.032)	−0.008 (−0.042, 0.025)
$\dot{\bar{a}}_{long}$	0.147** (0.032, 0.263)	0.221*** (0.094, 0.347)
$\dot{\bar{a}}_{lat}$	−0.087* (−0.182, 0.007)	−0.099** (−0.194, −0.005)
$present_{pass}$	−0.033 (−0.074, 0.007)	−0.033 (−0.074, 0.008)
$present_{lead}1{:}\bar{a}_{long}$		0.198** (0.045, 0.352)
$present_{lead}1{:}\dot{\bar{a}}_{long}$		−0.279** (−0.526, −0.032)
Constant	−0.043** (−0.082, −0.004)	−0.041** (−0.081, −0.002)
Observations	1,700	1,700
Log Likelihood	298.015	302.749
Akaike Inf. Crit.	−572.030	−577.498
Bayesian Inf. Crit.	−506.769	−501.360
Note:		*p<0.1; **p<0.05; ***p<0.01

Table B.4: Main and interaction effects for Maximum Heart Rate at the window level.

	Dependent variable:	
	Max. HR	
	b ($lower, upper$)	b ($lower, upper$)
t	−0.013 (−0.282, 0.257)	−0.028 (−0.301, 0.246)
$present_{lead}$	6.553*** (2.313, 10.792)	0.353 (−19.511, 20.216)
a_{long}^{max}	−1.954* (−4.208, 0.299)	−3.377** (−6.063, −0.690)
$dir_{a_{long}^{max}}$	−2.950* (−6.140, 0.240)	−2.121 (−5.456, 1.213)
a_{lat}^{max}	−13.713** (−25.268, −2.159)	−14.655** (−26.274, −3.035)
$dir_{a_{lat}^{max}}$	3.877* (−0.445, 8.200)	4.005* (−0.371, 8.381)
$\dot{a}_{long}^{max}$	−0.762 (−3.344, 1.820)	0.809 (−2.339, 3.957)
$dir_{\dot{a}_{long}^{max}}$	0.593 (−1.131, 2.316)	0.860 (−0.948, 2.668)
$\dot{a}_{lat}^{max}$	3.653 (−0.902, 8.209)	4.045* (−0.560, 8.651)
$dir_{\dot{a}_{lat}^{max}}$	−0.310 (−2.384, 1.764)	−0.442 (−2.512, 1.627)
$present_{pass}$	−1.345 (−4.803, 2.113)	−1.363 (−4.823, 2.097)
a_{long}^{max}:$dir_{a_{long}^{max}}$	5.554*** (1.401, 9.707)	7.201*** (2.508, 11.894)
a_{lat}^{max}:$dir_{a_{lat}^{max}}$	15.503** (3.679, 27.327)	16.379*** (4.505, 28.252)
$\dot{a}_{long}^{max}$:$dir_{\dot{a}_{long}^{max}}$	3.354 (−1.310, 8.018)	−0.401 (−5.969, 5.166)
$\dot{a}_{lat}^{max}$:$dir_{\dot{a}_{lat}^{max}}$	−7.821* (−16.037, 0.394)	−8.310** (−16.611, −0.008)
$present_{lead}1$:a_{long}^{max}		−0.339 (−7.089, 6.411)
$present_{lead}1$:$dir_{a_{long}^{max}}$		−10.009 (−36.294, 16.275)
$present_{lead}1$:$\dot{a}_{long}^{max}$		−2.574 (−8.036, 2.887)
$present_{lead}1$:$dir_{\dot{a}_{long}^{max}}$		−9.089 (−22.109, 3.930)
$present_{lead}1$:a_{long}^{max}:$dir_{a_{long}^{max}}$		6.783 (−4.241, 17.807)
$present_{lead}1$:$\dot{a}_{long}^{max}$:$dir_{\dot{a}_{long}^{max}}$		9.118** (1.042, 17.194)
Constant	109.982*** (101.093, 118.871)	109.774*** (100.855, 118.693)
Observations	1,700	1,700
Log Likelihood	−7,221.909	−7,215.102
Akaike Inf. Crit.	14,483.820	14,482.200
Bayesian Inf. Crit.	14,592.590	14,623.600
Note:		*p<0.1; **p<0.05; ***p<0.01

Table B.5: Main and interaction effects for Mean Heart Rate at the window level.

	Dependent variable:	
	Mean HR	
	b $(lower, upper)$	b $(lower, upper)$
t	−0.125 (−0.351, 0.101)	−0.145 (−0.374, 0.085)
$present_{lead}$	0.731 (−2.167, 3.629)	0.172 (−9.036, 9.381)
$\bar{a}_{long}$	0.946 (−2.655, 4.546)	1.696 (−2.222, 5.614)
$\bar{a}_{lat}$	1.442 (−0.901, 3.786)	1.613 (−0.751, 3.977)
$\bar{\dot{a}}_{long}$	−4.476 (−12.748, 3.796)	−6.678 (−15.730, 2.375)
$\bar{\dot{a}}_{lat}$	−3.170 (−9.784, 3.444)	−2.883 (−9.516, 3.749)
$present_{pass}$	−1.162 (−3.974, 1.649)	−1.218 (−4.033, 1.597)
$present_{lead}1{:}\bar{a}_{long}$		−2.959 (−13.503, 7.585)
$present_{lead}1{:}\bar{\dot{a}}_{long}$		9.781 (−7.181, 26.743)
Constant	106.595*** (99.705, 113.486)	106.644*** (99.749, 113.540)
Observations	1,700	1,700
Log Likelihood	−6,939.275	−6,938.579
Akaike Inf. Crit.	13,902.550	13,905.160
Bayesian Inf. Crit.	13,967.810	13,981.300
Note:		*p<0.1; **p<0.05; ***p<0.01

Table B.6: Main and interaction effects for Maximum Entropy at the window level.

	Dependent variable:	
	Max. Entropy	
	b $(lower, upper)$	b $(lower, upper)$
t	−0.0004 (−0.001, 0.0002)	−0.0004 (−0.001, 0.0002)
$present_{lead}$	0.014** (0.002, 0.027)	−0.042 (−0.107, 0.023)
a^{max}_{long}	0.002 (−0.005, 0.008)	0.006 (−0.002, 0.013)
$dir_{a^{max}_{long}}$	−0.004 (−0.012, 0.004)	−0.006 (−0.015, 0.003)
a^{max}_{lat}	−0.014 (−0.047, 0.019)	−0.010 (−0.043, 0.024)
$dir_{a^{max}_{lat}}$	0.011* (−0.001, 0.023)	0.009 (−0.004, 0.021)
$\bar{\dot{a}}_{long}$	−0.004 (−0.011, 0.004)	−0.004 (−0.013, 0.005)
$dir_{\dot{a}^{max}_{long}}$	0.001 (−0.005, 0.006)	−0.001 (−0.007, 0.005)
$\dot{a}^{max}_{lat}$	−0.003 (−0.017, 0.010)	−0.004 (−0.017, 0.010)
$dir_{\dot{a}^{max}_{lat}}$	0.0003 (−0.006, 0.007)	0.001 (−0.006, 0.007)
$present_{pass}$	0.002 (−0.007, 0.011)	0.002 (−0.007, 0.010)
a^{max}_{long}:$dir_{a^{max}_{long}}$	0.002 (−0.010, 0.013)	−0.004 (−0.017, 0.010)
a^{max}_{lat}:$dir_{a^{max}_{lat}}$	0.013 (−0.021, 0.046)	0.009 (−0.025, 0.043)
$\bar{\dot{a}}_{long}$:$dir_{\dot{a}^{max}_{long}}$	0.009 (−0.004, 0.022)	0.015* (−0.001, 0.030)
$\dot{a}^{max}_{lat}$:$dir_{\dot{a}^{max}_{lat}}$	0.011 (−0.012, 0.034)	0.011 (−0.012, 0.035)
$present_{lead}1$:a^{max}_{long}		−0.023** (−0.045, −0.001)
$present_{lead}1$:$dir_{a^{max}_{long}}$		0.034 (−0.046, 0.114)
$present_{lead}1$:$\bar{\dot{a}}_{long}$		−0.002 (−0.019, 0.016)
$present_{lead}1$:$dir_{\dot{a}^{max}_{long}}$		0.010 (−0.032, 0.051)
$present_{lead}1$:a^{max}_{long}:$dir_{a^{max}_{long}}$		0.032* (−0.0004, 0.065)
$present_{lead}1$:$\bar{\dot{a}}_{long}$:$dir_{\dot{a}^{max}_{long}}$		−0.008 (−0.033, 0.016)
Constant	0.223*** (0.202, 0.245)	0.227*** (0.205, 0.249)
Observations	1,352	1,352
Log Likelihood	2,338.157	2,341.406
Akaike Inf. Crit.	−4,638.315	−4,632.812
Bayesian Inf. Crit.	−4,539.337	−4,502.579
Note:		*p<0.1; **p<0.05; ***p<0.01

Table B.7: Main and interaction effects for Mean Entropy at the window level.

	Dependent variable:	
	Mean Entropy	
	b $(lower, upper)$	b $(lower, upper)$
t	−0.0003 (−0.001, 0.0002)	−0.0003 (−0.001, 0.0002)
$present_{lead}$	−0.004 (−0.015, 0.006)	−0.031* (−0.066, 0.004)
$\bar{a}_{long}$	−0.003 (−0.016, 0.011)	−0.004 (−0.018, 0.011)
$\bar{a}_{lat}$	0.002 (−0.006, 0.009)	0.002 (−0.006, 0.009)
$\dot{\bar{a}}_{long}$	0.012 (−0.014, 0.039)	0.010 (−0.019, 0.039)
$\dot{\bar{a}}_{lat}$	0.014 (−0.010, 0.038)	0.014 (−0.010, 0.038)
$present_{pass}$	0.004 (−0.005, 0.012)	0.003 (−0.005, 0.012)
$present_{lead}1{:}\bar{a}_{long}$		0.025 (−0.015, 0.065)
$present_{lead}1{:}\dot{\bar{a}}_{long}$		0.024 (−0.040, 0.087)
Constant	0.206*** (0.190, 0.222)	0.207*** (0.191, 0.223)
Observations	1,352	1,352
Log Likelihood	2,292.127	2,293.377
Akaike Inf. Crit.	−4,562.255	−4,560.755
Bayesian Inf. Crit.	−4,504.952	−4,493.034
Note:		*p<0.1; **p<0.05; ***p<0.01

Appendix C

Vehicle State Signals

Here, we present figures of all vehicle state signal samples. Each sample signal was obtained from a single trial.

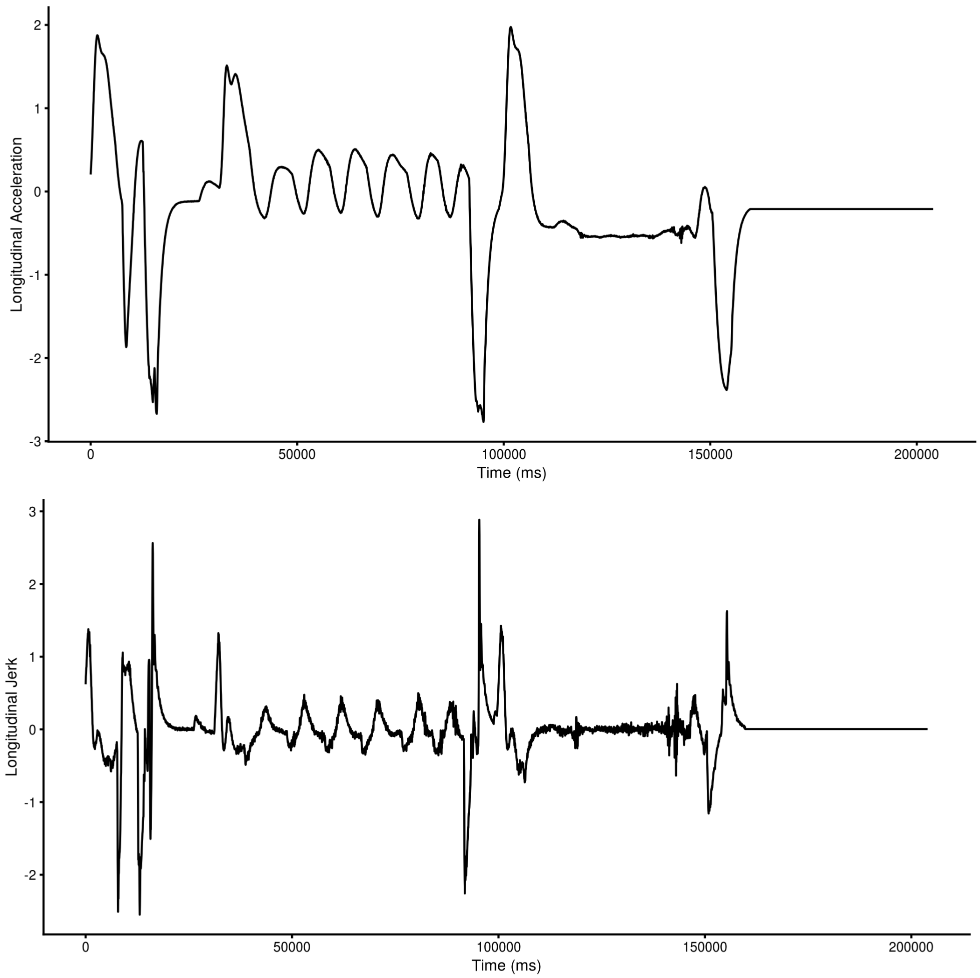


Figure C.1: Sample Longitudinal Acceleration and Jerk signals.

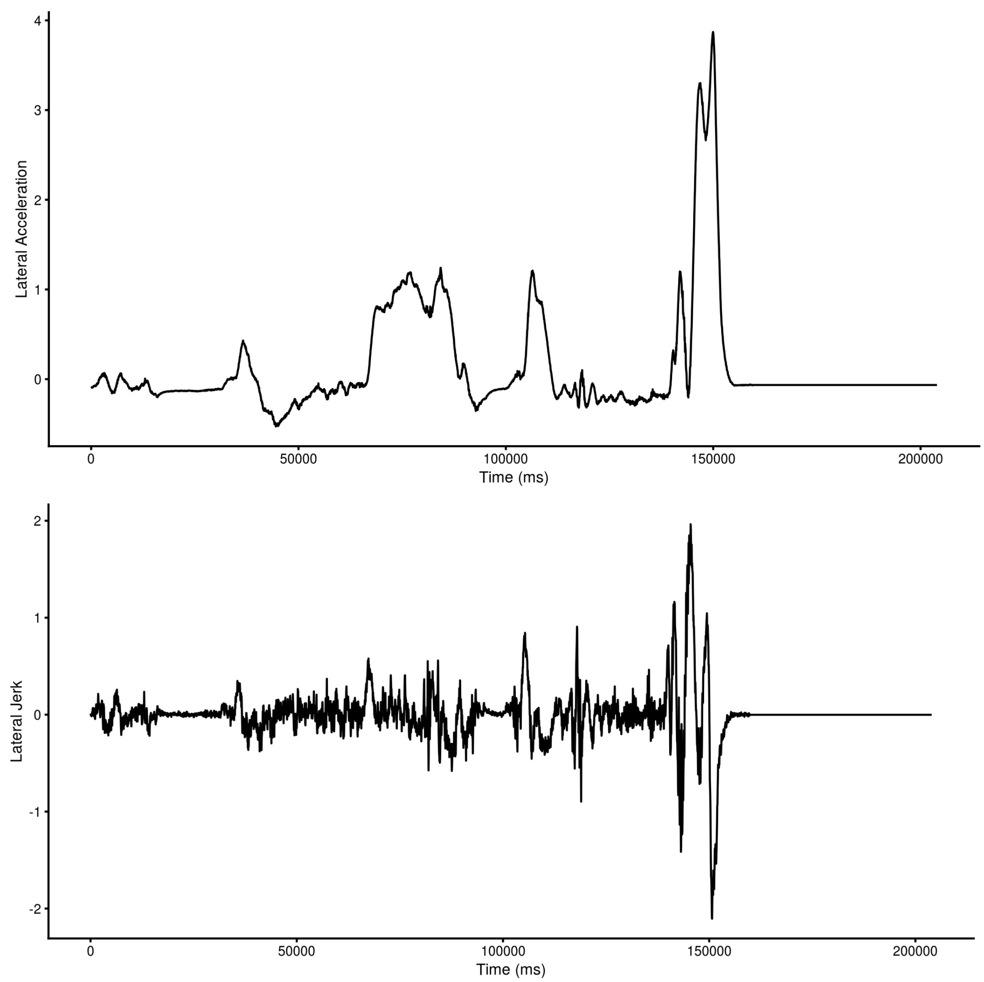


Figure C.2: Sample Lateral Acceleration and Jerk signals.

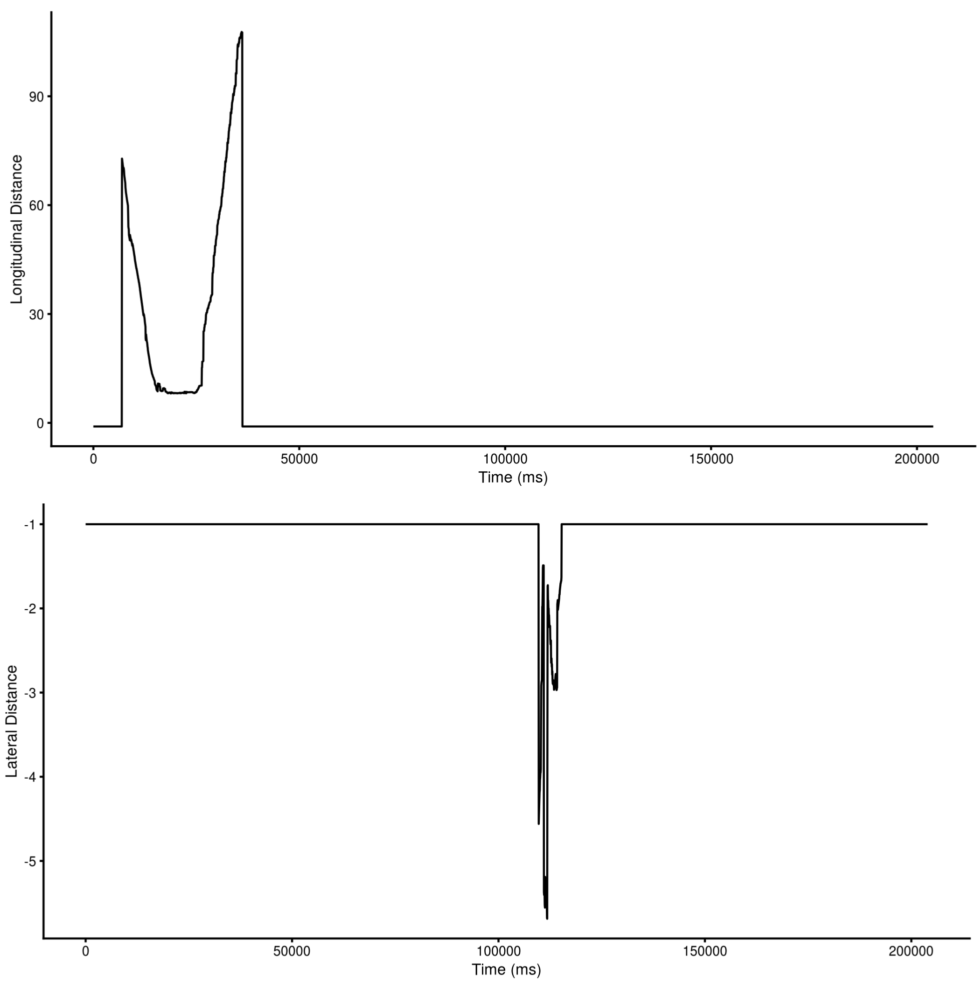


Figure C.3: Sample Longitudinal and Lateral distance signals. Values of -1 indicate the absence of a vehicle. Furthermore, lateral distances are negative as the agent vehicle was parked to the right of the ego.