DYNAMIC MODELING OF THE HUMAN-AUTONOMY COLLABORATION IN ADVANCED DRIVER-ASSISTANCE SYSTEMS (ADAS)

Kyle Moran

A Thesis Submitted to the
University of North Carolina Wilmington in Partial Fulfillment
of the Requirements for the Degree of
Master of Science

Department of Computer Science
Congdon School

University of North Carolina Wilmington

2024

Approved by

Advisory Committee

Lucas Layman	Geoffrey Stoker	Dale Cohen

Hamed Saeidi, Chair

Accepted By

Dean, Graduate School

TABLE OF CONTENTS

ABSTRACT

Dynamic Modeling of the Human-Autonomy Collaboration in Advanced Driver-Assistance Systems (ADAS). Moran, Kyle, 2024. Thesis. University of North Carolina Wilmington.

The rise of self-driving cars (SDCs) promises enhanced safety and convenience. However, these advancements introduce new challenges, particularly in the transitions between autonomous and manual control. Our research addresses these challenges by investigating human reactions during emergency takeover scenarios. To do this we build on previous research on the pitfalls of takeovers and handovers. Additionally, we employed the Psychological Value Theory (PVT) to analyze the effects of the perceived value of different decisions on the reactions of our participants. Our experimental design included dynamic scenarios that extend the static questionnaires used to study PVT in the past. Contrary to PVT's predictions, our results did not show a relationship between reaction time and psychological value but did exhibit a significant interaction between psychological value and participants' decisions on whom to sacrifice, as well as their gaze patterns during these decisions. These findings reinforce the applicability of PVT in dynamic environments.

Keywords - Self Driving Cars, Manual Takeover, Moral Dilemma, Dynamic Model, Mixed Control

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Acknowledgements

I would like to extend my deepest gratitude to all those who supported me throughout my graduate studies and the work on this thesis.

First and foremost, I am profoundly grateful to Professor Saeidi for his invaluable guidance and unwavering support. His patience, insightful feedback, and expertise have been critical to this project's success. By generously sharing his knowledge in problem-solving and new techniques, Professor Saeidi not only enhanced my ability to complete this work but also shaped me into a more well-rounded student.

I would also like to express my sincere thanks to Professor Cohen, whose work on Psychological Value Theory greatly influenced this research. His willingness to provide feedback on our experimental design, whether via email or during pilot testing sessions, was immensely helpful. His guidance enabled me to gain a deeper understanding of the subject.

A special thank you to Sirui Chen for his insights in our data analysis, and his dedication and willingness to work late into the night. His contributions accelerated our progress and has been instrumental in bringing this thesis to fruition.

Thank you to Kenneth Gatobu for his work on getting the project started and helping to initiate many of the tools used in this research.

I am also grateful to Professor Layman and Professor Stoker for their insightful discussions and valuable comments during the proposal stage. Their feedback helped to shape the direction of this project.

Lastly, I want to thank all those who supported me during my time as a student here, both faculty and staff. It is with their encouragement and assistance that I was able to reach this point, deepening my understanding of the subject matter and developing skills that will continue to inform my approach to learning throughout my life.

CHAPTER 1: INTRODUCTION

Advances in computer vision, autonomous control, machine learning, and artificial intelligence have enabled self-driving cars (SDC) with numerous autonomous features (Badue et al., 2021). The Society of Automotive Engineers (SAE (2021)) describes five levels of autonomy seen in Figure 1.

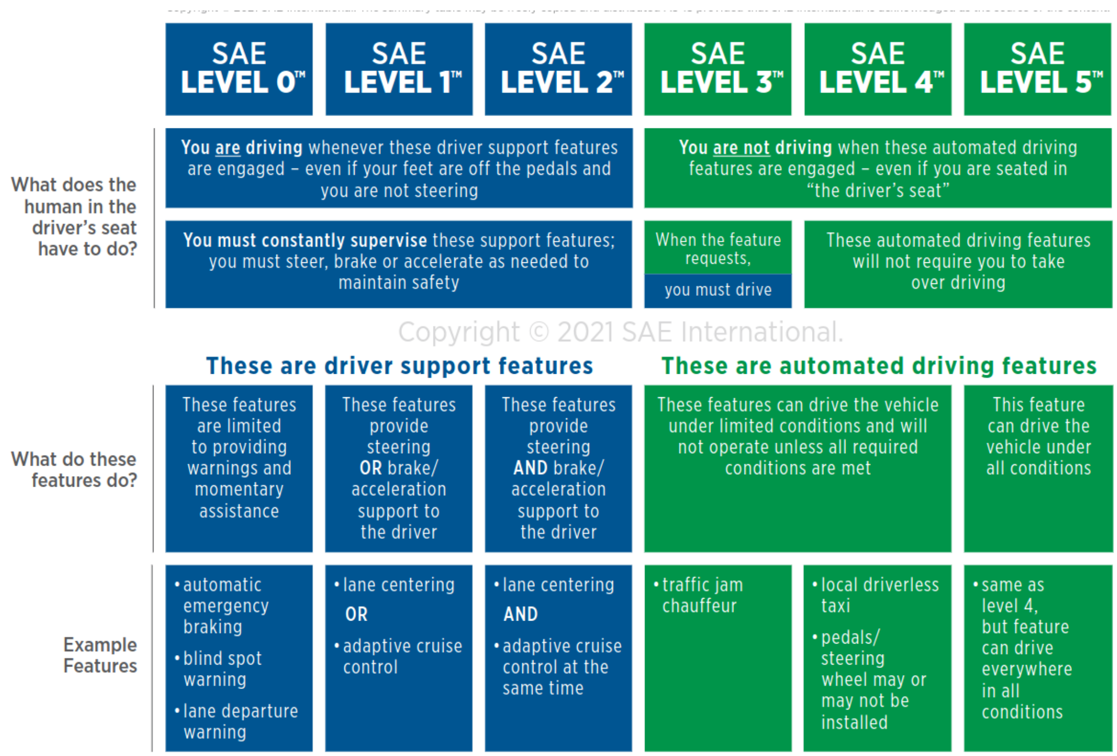


Figure 1. Graphic showing SAE's levels of automation (SAE International, 2021).

The majority of the existing self-driving cars only fulfill the criteria up to L3, though many companies have pilot programs based in a city where they are testing level 4 systems (Fox, 2021; Molla, 2021; Doll, 2023; Cozzens, 2022). Yet, there is no consensus timeline for achieving L5 due to the fluidity of the underlying technology and its liability implications for the parties involved. Some experts predict that L5 autonomy will not be achievable until 2075 (Shladover, 2016). Therefore, developing collaborative strategies between human drivers and autonomous features is crucial for safety and effectively utilizing state-of-the-art autonomous features while advancing towards full autonomy (Drexler et al., 2018, 2019).

Much of the push to use advanced driver assistance systems is for safety reasons. The National Highway Traffic Safety Administration (NHTSA) in the United States of America considers Autonomous Vehicles a primary factor in reducing the greatest cause of vehicular crashes: human error (NHTSA, 2016). Among these factors some are easy to correct for such as speed (Day et al., 2023). Others, like following directions and localization are quite difficult (Chakraborty et al., 2016). Despite some improvements, other new problems are created, like increased rates of distraction when monitoring a driver assistance system rather than driving a car completely by yourself (Baikejuli et al., 2023).

Improvements over the last few years have been quick and are likely to accelerate in the upcoming few. It is important to make sure that safety regulation and research keeps up with the improvements (NHTSA, 2016). The NTHSA provides twelve safety elements:

1. System Safety
2. Operation Design Domain
3. Object and Event Detection and Response
4. Fallback (Minimal Risk Condition)
5. Validation Methods
6. Human Machine Interface
7. Vehicle Cybersecurity
8. Crashworthiness
9. Post-Crash ADS Behavior

10. Data Recording

11. Consumer Education and Training

12. Federal, State, and Local Laws

Our work focuses primarily on the fallback and possible transition from autonomous to manual. This process in vehicles lower than L4 is discussed in Brandenburg (2021). They lay out four stages of an ADAS (see Figure 2). States (1) and (3), manual and autonomous respectively, are continuous states in which the system is at equilibrium. States (2) and (4), handover and takeover, define the transient behavior of the system during the control mode switches. These states pose several challenges regarding trust, confidence, and situational awareness issues.

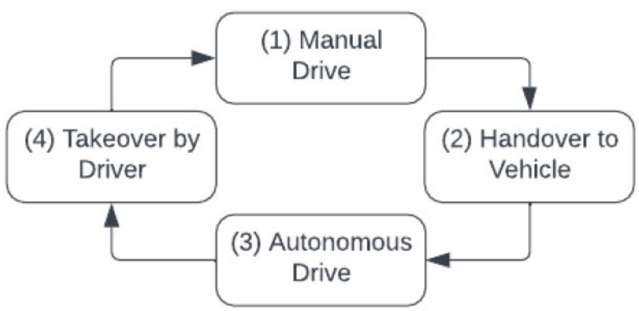


Figure 2. Four stages of semi-autonomous driving (Brandenburg, 2021).

The handover problem presents unique human interaction questions. If the driver does not trust the autonomous system and refuses to enter the handover state, the benefits gained by having an autonomous system will not be achievable. Despite myriad safety and productivity gains (NHTSA, 2016), usage of this technology is not a given, especially in the elderly (Lee et al., 2019). On top of the high-level goal of increasing trust in ADAS, another equally important pursuit in improving the handover stage is decreasing the perceived load and making the transition as seamless as possible (Zhang et al., 2020).

Similar, but somewhat opposing, issues exist in the takeover state. Inevitably an ADAS will run into scenarios in which it does not have the ability or confidence to continue driving safely; in these scenarios, the driver will have to take over vehicle operations. A driver who does not have to control the car is likely to have reduced alertness and will take longer to react (Banks et al., 2018). In these situations, a takeover request may come as a surprise to a driver who has “fallen asleep at the wheel”. Overcorrection, panic, and slow response times are common for those who have stopped paying attention to what the car is doing. Improving this step in the automated driving loop is important to the relationship between drivers and the systems in their cars (Epple et al., 2018).

In this work, we focus on modeling the dynamic behavior of drivers during state (4). The goal is to use these models to predict the outcomes of different control modes and the transitions to enable mode switches that are intuitive for human drivers. Furthermore, these models predict whether control mode switches can improve the overall performance of the system as compared to single-mode control. We consider the cases when a crash is seemingly inevitable or an averted one becomes dangerous to the operator. We must ask questions of how to “best” proceed with the accident. This can be framed as an ethical dilemma, or more specifically, a *sacrificial moral dilemma* (Goodall, 2014; Bartels and Pizarro, 2011). This refers to the situations in which, regardless of the choices made by the actor, death will be a result. In discussions surrounding sacrificial moral dilemmas, choices are often described as either utilitarian or deontological. Consider the crying baby dilemma:

> [either] smother one’s crying baby to death in order to prevent its crying from summoning enemy soldiers who will kill oneself, the baby, and a number of others if summoned (Greene et al., 2001).

The utilitarian choice would be the one that maximizes the values of the outcome. In the context of the above scenario, this means smothering the child. On the other hand, the deontological choice would be harm reduction. A deontologist, conversely, might not smother the child (Christov-Moore et al., 2017). In order to analyze a participant's choices, we will utilize the Psychological Value Theory which explains how humans make decisions in such scenarios based on the perceived values of the decision criteria (Cohen and Ahn, 2016).

In this work, we developed an experimental setup to identify how this theory affects the behavior of the drivers in a dynamic scenario when they are collaborating with autonomous car controllers. We track the attention of the driver (via gaze-tracking), their reactions during the test cases (how the steering wheel and pedal controls are modified during the take-over step), what decisions they made (what pedestrians to hit), and the consequences of such dynamics on the outcomes of shared driving scenarios.

The remainder of this thesis is organized according to the following: A detailed literature review of the problem is presented in Chapter 2. The technical details of the experimental setup and the corresponding hypothesis and measured variables are explained in Chapter 3. The results from the data gathered from our experimentation are presented in Chapter 4 and discussed further in Chapter 5. Finally, we will speak to the next steps for this project and how our findings can be improved upon and extended in the future in Chapter 6.

CHAPTER 2: LITERATURE REVIEW

2.1 Self-Driving Architecture

Autonomous controllers in SDCs take various forms. A level 1 autonomous vehicle (AV) uses the current car speed and the distance to the car in front to realize a closed-loop adaptive cruise control. Typically, in the feedback loop, a radar sensor in combination with the speedometer will feedback to a controller which autonomously updates the car's throttle and/or brake in order to maintain a safe distance and velocity on the road. The result is comparable to, and in some cases better than, its human counterpart. In a level 2 system, several other sensors and control algorithms are used to assist with staying in the lane in addition to the L1 features. Usually, several RGB cameras are used around the car to detect lane lines and the vehicle's position inside them. Other sensors gather data about the current steering angle and steering force. More intricate controllers coordinate the overall motion of the car in such scenarios (controlling the throttle, brake, and steering wheel) (Zanchin et al., 2017).

Level 3 autonomy is the highest level where a human is still in the loop, but they are much further removed than in level 2. The most popular example of a level 3 controller is traffic jam assist (Magazine, 2017; SAE International, 2021). When serving in this role, no human input is needed until the vehicle exits the environment in which it can comfortably operate. It will then resume mixed control with the operator. A vehicle at this level of automation does not have route planning abilities like those of a level 4 vehicle but does need the ability to completely monitor its surroundings. Level 3 vehicles require not only an increase in the kinds of sensors but more importantly an increase in quantity, as the level 3 vehicle needs to be aware of everything happening near it (Zanchin et al., 2017).

Level 4 autonomy enables the SDC to make simple, short-term driving decisions, as well as the ability to plan long-term and adjust that plan based on its short-term decisions. This adds a route planning system to the complexities of the sensors (see Figure 3) (Badue et al., 2021). The jump from level 3 to level 4 includes some new hardware like pedestrian-sensing lidar, but most of the gains are made via sensor fusion, which is the ability of controllers to combine data from several sensors in a robust ecosystem (Chakraborty et al., 2016). For example, in path planning this RGB cameras update the local area map which make GPS coordinates accurate to the inch (Badue et al., 2021). Sensor fusion is also used in developing a fail-safe strategy should a sensor on the car fail (e.g., if debris breaks a side RGB camera, a lidar detector can still "see" obstacles in the blind spot). This decreases the need for constant monitoring (Zanchin et al., 2017).

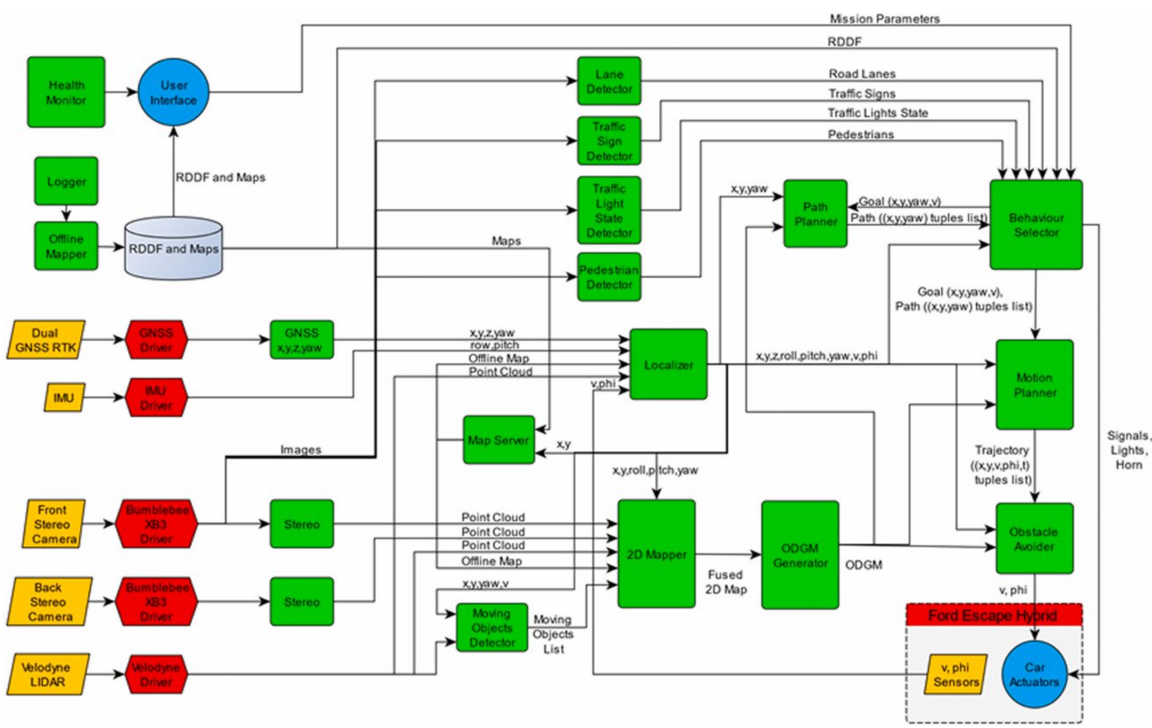


Figure 3. Block diagram of all required sensors and decision-making systems in a level 4+ autonomous vehicle (Badue et al., 2021).

While several cutting-edge technologies are being developed and tested for SDCs, the majority of commercially available cars are level 2 and below. Hence, the vast majority of driving tasks are still handled by a human (Counterpoint, 2022). Human drivers are responsible for an average of over 100 deaths per day (NHTSA, 2023). Autonomous driving technologies, such as collision prevention, have made cars generally safer (NHTSA, 2016). However, as long as there is some manual input to the control loop (i.e. levels 0-3), the human tendency to make mistakes will continue to cause safety concerns. Particularly during monitoring tasks, operators of an ADAS exhibit riskier behavior than when manually controlling the car (Banks et al., 2018).

Like humans, autonomous controllers are fallible given the limitations in sensing, planning, and control technologies — especially in complex and dynamic real-world scenarios. Even the autopilot systems of planes fail despite being used in a less complicated environment (Beringer et al., 1997). In these cases, the pilots are often considered negligent as they, much like the ADAS operator, are in charge of monitoring (U.S., 2012). Therefore, there is a clear need for dynamic control mode changes during the operations of vehicles.

2.2 Takeover and Handover Strategies

The time it takes for a driver to press the brake when an obstacle becomes visible is called the perception-response time (PRT) (AASHTO, 2011). The first part of this reaction, the perception, is the amount of time it takes to process what you are seeing. Once the driver recognizes that there is an obstacle, they need to send the signal to the foot to apply the brake (See Table 1 for the standard numerical values for PRT). These are calculated for a "surprised," attentive driver. If the driver is expecting to get into an accident or if they are distracted while driving these times change (Wood and Zhang, 2021; Lieu and Koppa, 1999).

Table 1. Perception Response Times*.

Component	Individual Time (sec)	Cumulative Time (sec)
Latency	0.31	0.31
Eye Movement	0.09	0.40
Fixation	0.20	1.00
Recognition	0.50	1.50
Brake Application	1.24	2.74

*The top four components are perception. The bottom is response (Lieu and Koppa, 1999)

In a meta-analysis of 36 studies, Soares et al. (2021) corroborate the findings of the United States Department of Transportation (DOT). Most (75%) of the studies had PRTs between 1.5s and 3.5s. Experiments with more complicated design, especially those which introduce distractions, tend to increase PRT and the risk of collision. Conversely, simple tasks that focus only on operation had lower PRTs. These results only measure the time it takes to begin a reaction rather than the quality of that reaction. Measuring quality is more challenging. Generally, a faster PRT does result in a safer takeover (Soares et al., 2021; Eriksson and Stanton, 2017).

Accurately predicting how a driver will react allows for a shared control scheme that gives the driver more time to react. In order to accomplish this, Jin et al. (2021) employ a modified ACT-R model (Anderson, 2007) to outline the characteristics that would determine the “success” of the takeover (see Figure 4). The researchers then measure the influence of each factor on the result, both primary (e.g., traffic density) and secondary (e.g., trust).

From these results, Jin et al. created a Structural Equation Model (SEM) that forecasts the takeover’s success (see Figure 4). Success combines variables: takeover time and quality. The former is the amount of time until control is restored, and the latter, is a combination of several factors, including the longitudinal/horizontal acceleration and deviation from the center of the lane. By building on the highest weighted factors, we can use their findings to offer a valuable framework to predict and reinforce a safe and speedy takeover loop.

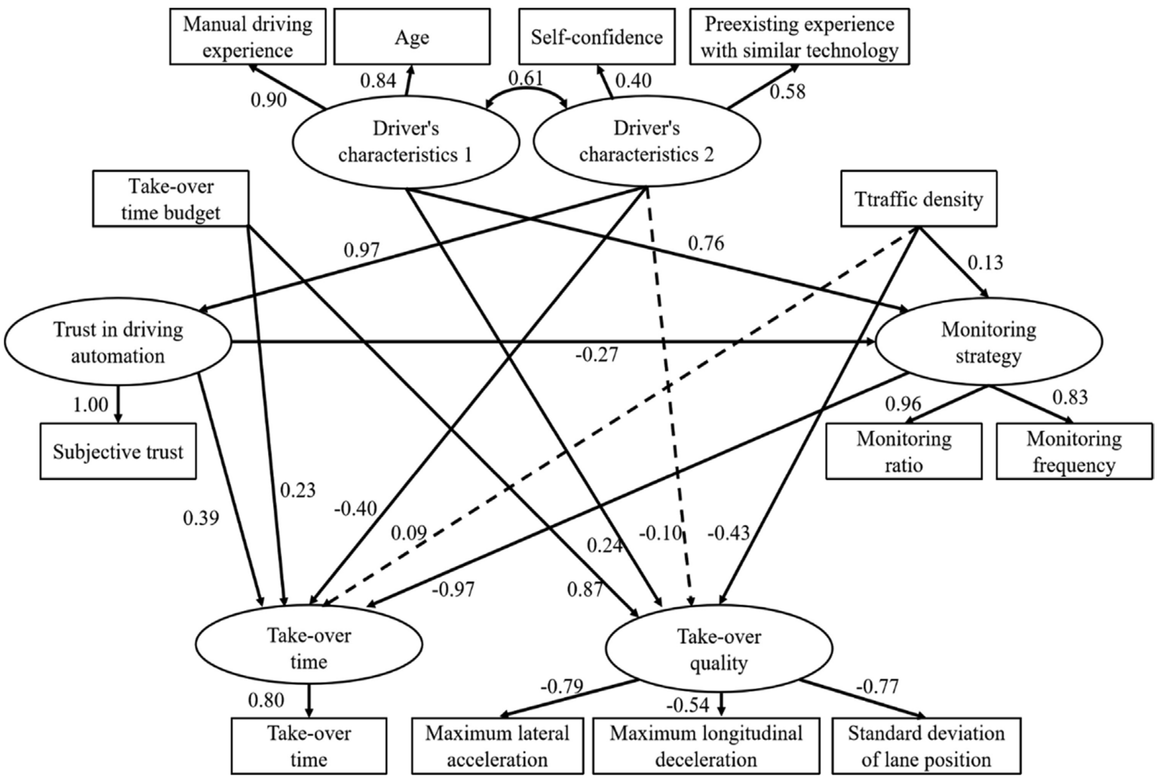


Figure 4. Jin et al.'s final SEM with normalized path coefficients (Jin et al., 2021).

2.3 Failures

Failures in autonomous systems can be split into two groups: silent loud failures (Gold et al., 2013; Fu et al., 2020).

Loud failures alert the user when the autonomous system is going to fail or disengage. These are strongly preferred for many reasons, including improving alertness and overcorrection Zhang et al. (2023). Additionally, loud failures are one of the fallback techniques suggested in the NHTSA safety elements (NHTSA, 2016). The implementation of loud failures is a challenge because the autonomous system must be able to recognize something is wrong in order to alert the operator for them to takeover.

Conversely, silent failures occur when an autonomous system is not aware anything is wrong and can be more dangerous as the operator is not aware they need to take over. Silent failures have negative effects on trust (Fu et al., 2020), which can, as reduce the trust of the driver to handover control and limit safety gains (Drexler et al., 2018).

Much of the current research is focused on improving loud failure techniques since they are the key to making autonomous vehicles more safe for the operator. To avoid these problems, some researchers propose the inclusion of recurring alerts that ensure an operator is paying attention. This can be achieved either by monitoring the user or at fixed intervals. In a similar experiment, Merat et al. (2014) found that fixed intervals were more successful in keeping the driver alert rather than monitoring their eye movements.

Our research focuses primarily on reactions in silent failures and how the operator is able to comprehend and make decisions when a silent failure occurs. In general silent failures lead to a higher PRT and more variable reactions. These combine to be less safe than loud failures (Bianchi Piccinini et al., 2020; Mole et al., 2020).

2.4 Gaze When Operating an Autonomous Vehicle

Gaze information is often used as a way of monitoring the driver to ensure they are paying attention (Merat et al., 2014; Hofbauer et al., 2020). Through the use of what they termed *visual gist*, Pugeault and Bowden (2015) were able to infer the driving context with a reliability of ≈80% up to a second before the driver completed the task. *Visual gist* a quick survey of what is taking place of a scene and how the eyes take it into account. The Fidelity of *visual gist* did decrease in more complex road scenarios. When testing to see if including peripheral gaze improved their predictions, Pugeault and Bowden (2015) found no significant improvement unless at a crossroad where looking to both sides for oncoming traffic is required.

In an effort to show a more robust use for gaze, Fan et al. (2021) combine exiting gaze frameworks from a simple RGB camera and head position to predict what action a driver is going to take as well as if they are paying attention (Figure 5).

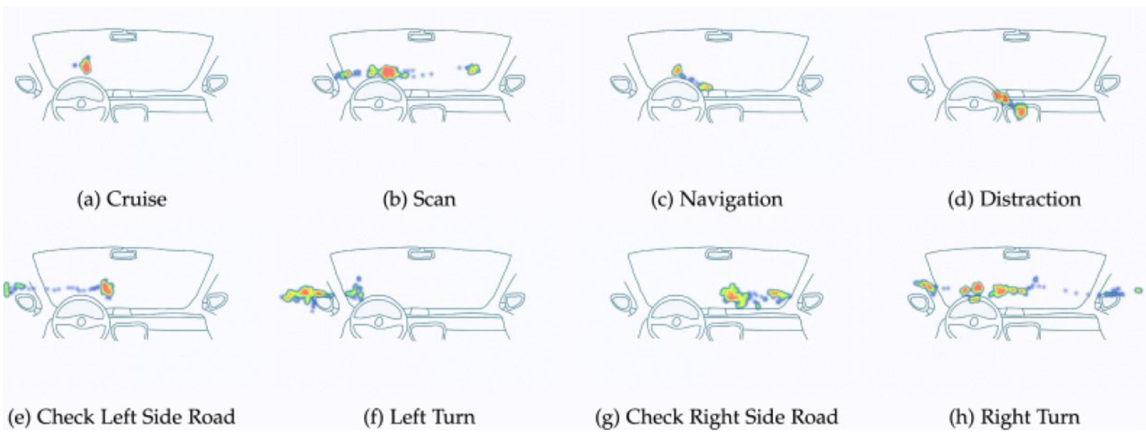


Figure 5. Example of gaze heatmap from Fan et al. (2021) showing how different patterns appear in different driving maneuvers.

2.5 Psychological Models of Human Intervention

Most of us do not often have to make life-or-death decisions. Driving, however, is an exception. Generally, we trust people to make decisions based on their own moral framework. But when machines will be making decisions instead of humans, we need to decide how it will make ethical decisions. When questioned about a car in a sacrificial moral dilemma, most respondents will elect for the machine to make the utilitarian choice. However, when asked if they would purchase a vehicle that makes utilitarian decisions, the responses shifted overwhelmingly to "no" (Bonnefon et al., 2016; Gill, 2020). Because humans do not consider all lives as equally valuable (e.g., saving your mother is preferential to saving a stranger) (Cohen et al., 2022), consumers are not likely to purchase autonomous driving systems that make purely utilitarian decisions. The discrepancy in perceived value becomes especially apparent when passengers are added to the above situations. When asked if a self-driving car should sacrifice a family member or five strangers, respondents said the car should sacrifice the strangers (Bonnefon et al., 2016).

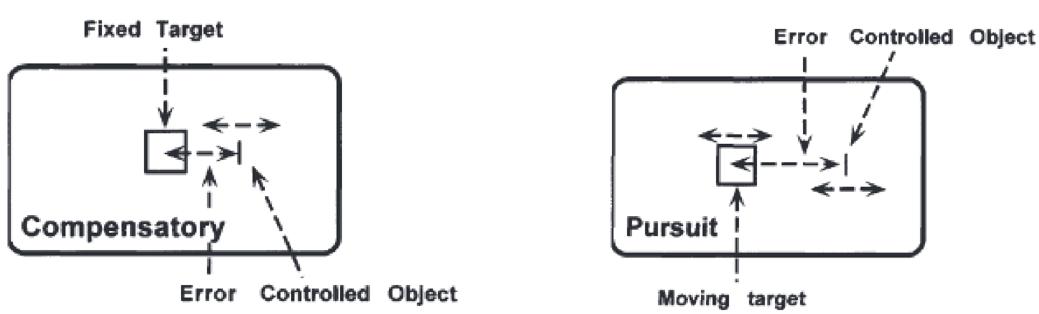
In addition to finding a discrepancy between the value of certain human lives, Cohen et al. (2022) also observed a link between the psychological value and PRT. That is, as the distribution of the perceived values of the two choices converge, the PRT for choosing a sacrifice increases. The researchers showed that when making a decision, lives, like economic goods, are assigned a value. Importantly, these decisions seem to be non-linearly affected by group size. For example, a participant might assign twice the value to a judge as compared to a nun. So, when faced with a decision of whether to sacrifice four nuns or a judge, they would likely choose the latter. This runs counter to the expectation of utilitarian decision making.

2.6 Quantitative Approaches to Modeling Humans as Controllers

Humans do not sample continuously. Instead, we make discrete checks at various time intervals and adjust control accordingly. This is called asynchronous control (Jagacinski and Flach, 2003a). When monitoring a self-driving car, humans have a lower sample frequency versus when they drive (SWOV, 2012). This is particularly dangerous when allowing an ADAS to control the car. If an obstacle approaches quickly and an emergency takeover request is required, it may take too long to increase the sample rate and recover safely. In order to combat this, we can use sensors that continually monitor the environment to supplement human control and create a safer and more respondent system (Badue et al., 2021).

In many cases, a human's operation of a car resembles that of a "fuzzy controller". A fuzzy controller consists of multiple control states any of which may be turned on at once. Sugeno and Murakami (1984) displayed the ability to model a human's behavior using this control strategy. They asked participants to park a model car, and after observing its movements, developed 42 fuzzy rules that achieve results similar to those of a human.

Such a control system is useful for tasks like parking, where the goal is to maneuver a visible point to stay in a certain position (e.g., the center of a parking spot to the center of the car). This is a *compensatory task* (see Figure 6a). Avoiding a stationary pedestrian is still compensatory, but instead of trying to keep the frame on the target, moves it away from the target. When the pedestrian is moving, the task becomes avoidant *pursuit* (see Figure 6b). By defining the problem in these terms, the dynamic system is stabilized by optimizing the mean squared error (MSE) (Jagacinski and Flach, 2003b). In a normal tracking task, optimization is accomplished by minimizing the MSE; however, our goal is to track to avoid the pedestrian, so we want to maximize the pursuit MSE. Simultaneously, we need to minimize the MSE of the compensatory task of staying on the road.



(a) Example of a compensatory task. (b) Example of a pursuit task.

Figure 6. Example of both types of types of tracking tasks (Jagacinski and Flach, 2003b).

CHAPTER 3: METHODOLOGY

3.1 Problem Statement

With the goal of creating better shared control strategies for operation of a self-driving car, we needed to understand how the operator would react in high-stress or emergency situations. It is with this goal in mind that we approached this research.

In order to create a simple emergency scenario, we collaborated with, and based much of our research on, the work done by Dr. Cohen on PVT (Cohen and Ahn, 2016). When studying PVT, Cohen used static questionnaires (e.g., Figure 7) to empirically test PVT's rigor. Here, we aim to extend these questionnaires to dynamic scenarios. Building on top of PVT gave us previously computed reaction times as well as a decision making framework from which we could model the reaction of our participants.

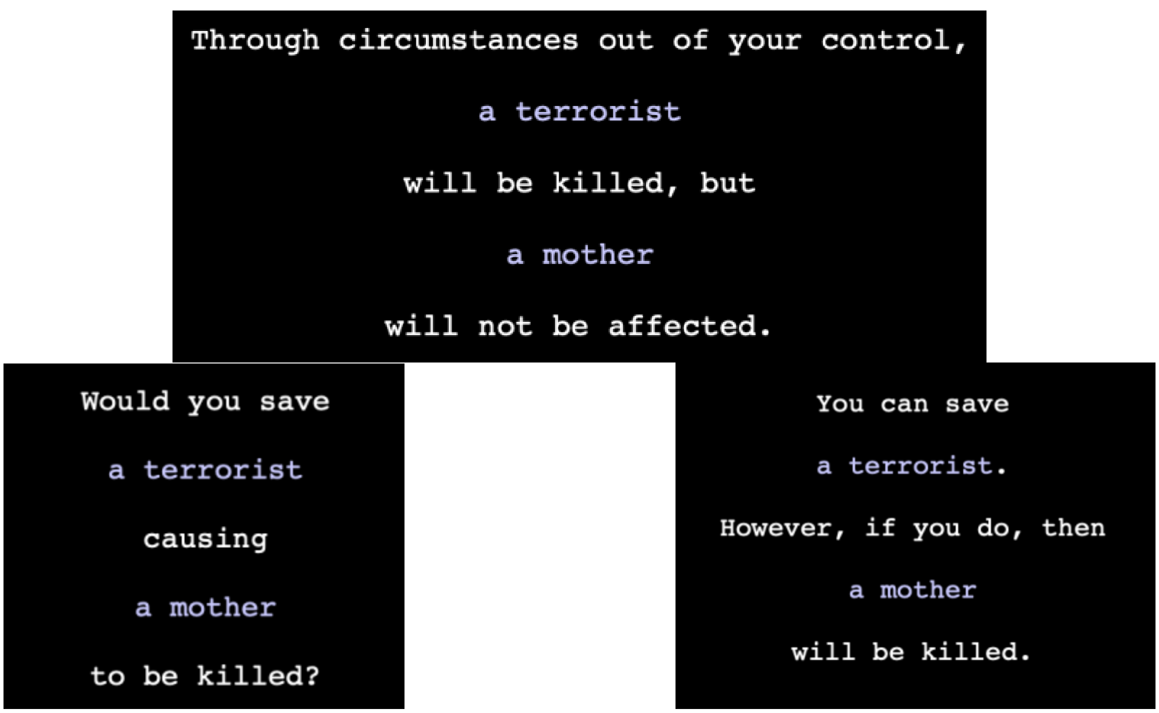


Figure 7. Example of the questionnaire used in Cohen et al. (2022).

As previously established, the takeover in emergency scenarios is an important area of study because humans in emergencies tend to overcorrect and the autonomous driving systems that currently exist are not expected to succeed all the time (Brown and Laurier, 2017).

PVT leads us to expect that the following will be true:

- The reaction time is negatively correlated with the difference in overlap of a pedestrian's psychological value ($|\Delta_v|$).
- Drivers choose to collide with the lower value target the majority of time and do so more consistently when the $|\Delta_v|$ is higher.

Once we have the results from our experimentation, we can assess the accuracy of our hypotheses and use the data to construct dynamic models of real-time human decision-making patterns. The success of such models will be measured by how well they can predict the actions of the operators.

3.2 Scenario Design

As mentioned in the previous section, the primary influence on our line of experimentation is the work in Cohen and Ahn (2016). In order to faithfully convert the questionnaires into simulations we tried to biject each element in the questionnaire to one in the simulator (see Figure 8). Clearly, some of the adjustments needed (e.g., labels that describe the pedestrians) do not reflect real life. These were made as a concession in order to still be able to test the applicability of PVT in our simulation use case.

3.2.1 Choosing a Sacrifice. PVT makes it clear that categories of people have a similar value when judged by the majority of the population. Ensuring these values hold in our simulation was important when designing the scenarios. We settled on the admittedly simple solution of labelling the pedestrian in a similar vein to the original PVT experiment. We then needed to choose what labels to show in order to limit the number of variables and to be able to test the relationship between values. We decided to only include a representative sample of labels from which we can draw conclusions about the applicability of PVT rather than its efficacy.

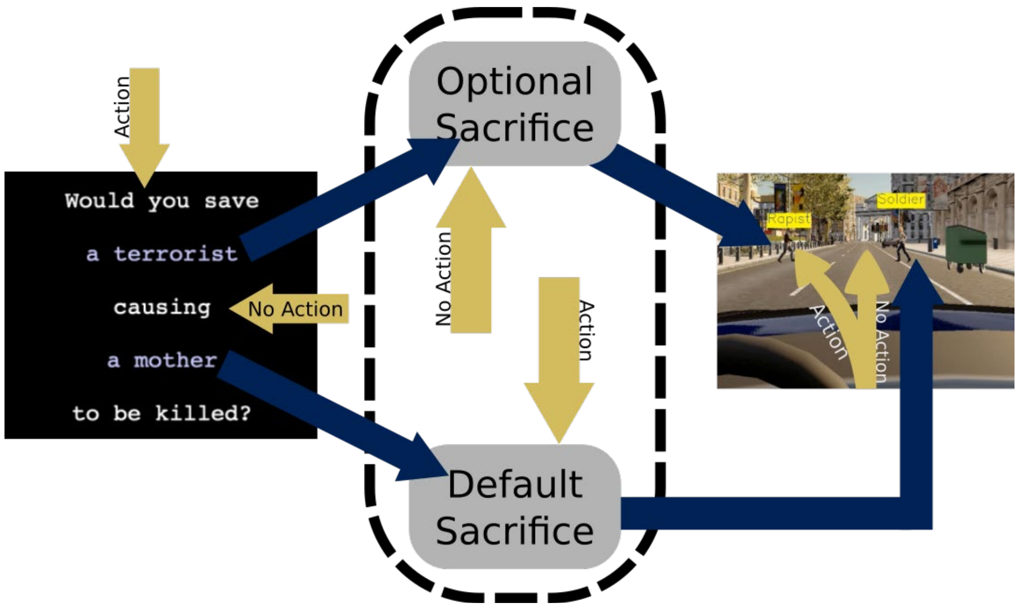
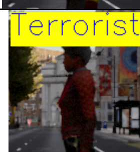
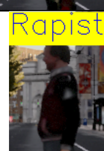
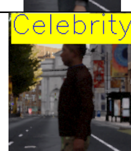
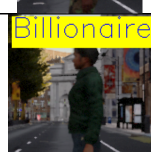
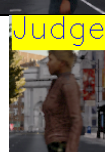
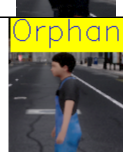
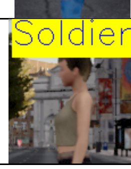


Figure 8. One-to-one mapping from the questionnaire to the simulation.
The labels are preserved as well as the necessity to act in order to save one pedestrian and sacrifice the other.

We determined that we could split the values of pedestrians into three broad categories delineated by value as described in Cohen et al. (2022) Table 2: Low, Medium, and High. These three categories would be used to determine overlap in the same way that a single object is used in the static trials. The advantage of this is that it allows us to keep the number of test conditions low so that participants can finish the experiment in a 1-hour session while providing us with a varied look into the effects of different numerical overlaps.

In order to determine the actual model of the pedestrians used in the simulator we randomly chose pre-built models from CARLA's (CAR Learning to Act) blueprint library (Dosovitskiy et al., 2017). However, in the case of the orphan and the police officer we used models that could be more easily recognized as a police officer and as a child (see Police and Orphan in Table 2).

Table 2. Images of the Models with Their Labels and Value Distribution for Each.*

		Psychological Value		
Label	Image	25%	50%	75%
Pedophile		0.0	0.0	11.25
Terrorist		0.0	0.0	100.0
Rapist		0.0	0.0	100.0
Celebrity		1,000.0	3,000.0	20,000.0
Billionaire		1,000.0	4,000.0	50,000.0
Judge		1,000.0	5,000.0	10,000.0
Police		2,000.0	10,000.0	85,000.0
Orphan		4,000.0	10,000.0	100,000.0
Soldier		6,000.0	15,000.0	1,000,000.0

**“[The] values for the response quartiles represent the raw responses compared with the standard, chimpanzee = 1,000 (Cohen et al., 2022)”

3.2.2 Psychological Value. PVT has shown that a more difficult choice (i.e. a choice with a low Δ_v) increases reaction time and also increase the chance of an incorrect choice (Cohen et al., 2022). Of our trials, 1/3 were high overlap (denoted as Δ_v=0 here), and 1/3 of those were full overlap, meaning that the same value target appears on both sides of the street. In contrast, in trials where the relative value of the pedestrians was moderately close, we denote this as $|\Delta_v| = 1$. Values with very little overlap are $|\Delta_v| = 2$. $\Delta_v = 0$ served as our control when balancing the default reaction when a pedestrian appeared. For example, when conducting pilot experiments, we found that pilot testers tended to stay in their own lane rather than change lanes. Cases with full overlap help us to identify bias. In later models we can remove this bias to attain better results for operator preference. There will be more discussion on this effect in section 3.3.5.

3.3 Implementation of the Scenarios

In order to assess the dynamic behavior of participants in a closed environment we used a driving simulator. The simulator, CARLA (Dosovitskiy et al., 2017) v0.9.14, built using Unreal Engine 4.26 (Epic Games, 2019), is used for simulating highly automated vehicles in various environments (Hofbauer et al., 2020; Santos and Larocca, 2019) (see Figure 9). In order to build these scenarios, we controlled the environment and the autonomous controller through the use of Robot Operating System 2 (ROS2) Foxy Fitzroy. It enables a standard robotic system application development framework via several tools, giving us the capability of measuring and controlling without any latency from one affecting the other (Macenski et al., 2022). The CARLA repository provides a bridge that serves as an interface between ROS and CARLA (Carla Simulator, 2022).

3.3.1 Autonomous Control. By default, the bridge contains basic autonomous controls like stopping at stoplights and cruise control. The autonomous control is split into 3 parts: a local planner, longitudinal controller (Assisted Cruise Control and Object Detection), and lateral controller (Lane Keep Assist) (See Figure 10). The local planner uses the ground truth map in CARLA to generate a path along the road of waypoints (Figure 11). Waypoints are pseudo "GPS" coordinates in the world frame located in the center of the lane.

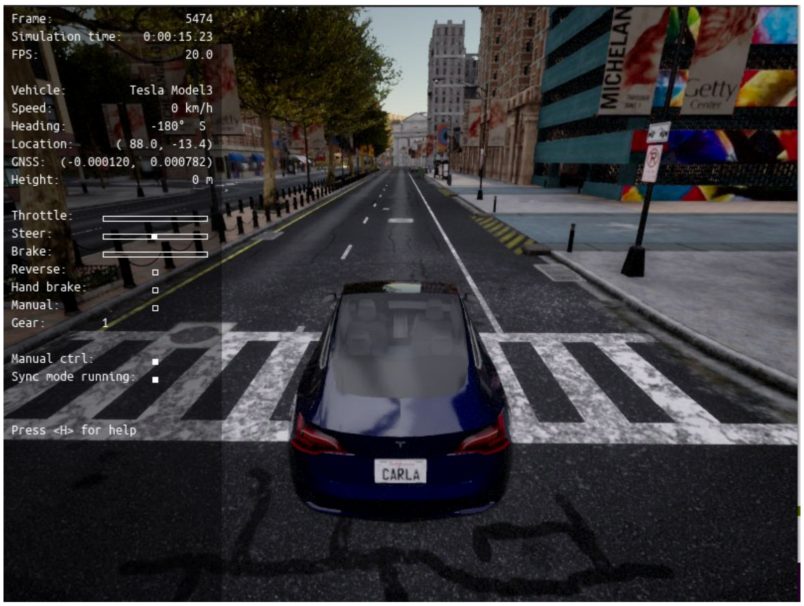


Figure 9. Example of car driving in the CARLA simulator.

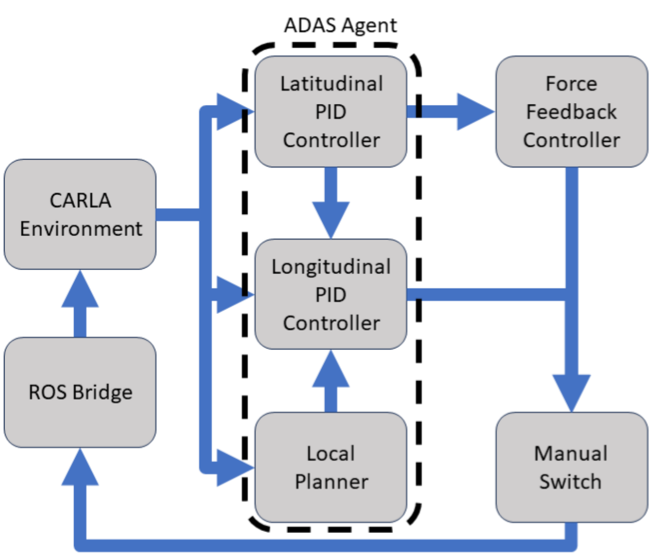


Figure 10. Control structure of how the ADAS agent in our experiments is fed data and how controllers interact.

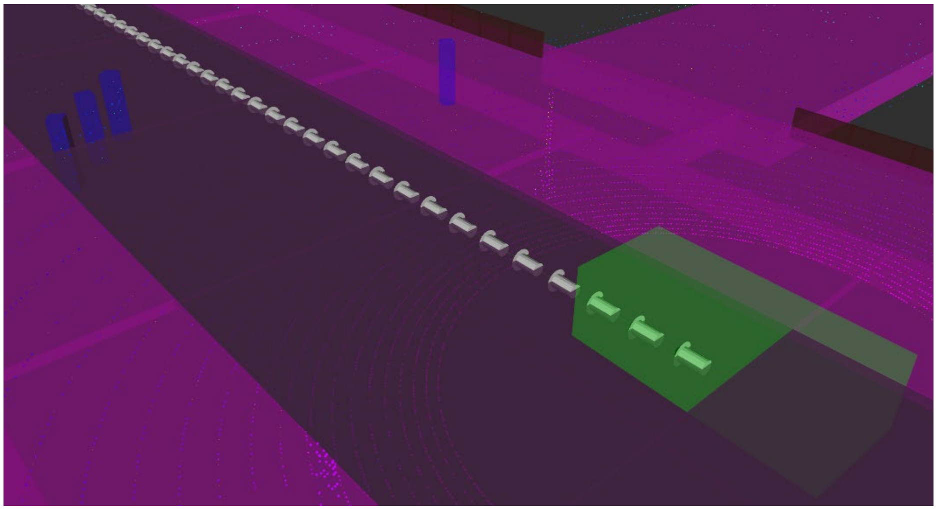


Figure 11. Image showing the waypoints in CARLA and a pedestrian nearing them.

Both of the other controllers, longitudinal and lateral, make use of proportional integral derivative (PID) controls. Each component of a PID controller has its own role. The P-component proportionally increases the control based on the error, the I-component reduces steady-state error over time, and the D-component dampens the output. Equation (1) is used for calculating the output of a PID controller (Samak et al., 2020). The coefficients k_p, k_I, k_D are tuned by hand in order to receive the desired output of the controller.

$$u(t) = k_P * e(t) + k_I * \int_{t_0}^{t} e(t)dt + k_D * \frac{d}{dt}e(t) \tag{1}$$

The lateral controller uses the waypoints to make sure that the vehicle remains in the center of the lane. It does this by staying as close as possible to the next upcoming waypoint. If the vehicle is not directly over the waypoint, the PID controller will use the distance away from the waypoint as the error and the output will go to the steering command.

The cruise control part of the longitudinal controller takes in a target speed — determined by the speed limit of the road — and uses the PID controller to modulate the throttle until the target speed is reached.

The last element of the autonomous controller is object detection. Detecting whether something is in the lane is as simple as checking if any obstacles share a location with any of the waypoints and then setting a slower target speed to avoid a collision. The difficulty comes in detecting a pedestrian's intention to cross the street. Predicting intention requires measuring the pedestrian's angular velocity and yaw in order to assess whether or not they are going to continue on the sidewalk or enter the road, in which case we would either slow down to a safe speed or stop the vehicle entirely.

3.3.2 Manual Control. Designing manual control required more focus on interfacing with the driver rather than the controllers in the car itself. Most of the interfacing was done through the Logitech G29 controller (Figure 12) and a ROS compatible driver (Atsushi, 2023).

(a) A Logitech G29 Racing Wheel. The buttons labelled "A" are the paddle shifters. These are used to turn on autonomous

(b) Logitech g29 Racing pedals. Only the two labelled pedals are used, the gas and brake

Figure 12: Logitech G29 Racing wheel with pedals.

The CARLA to ROS bridge includes a manual controller which allows for a keyboard to operate the vehicle. We extended that functionality to include operation by the steering wheel. In order to read manual inputs from the steering wheel, we modified a driver designed for controlling the G29 steering wheel in ROS and its ability to provide force feedback (Atsushi, 2023). In order to make the simulation more realistic, we use force feedback to automatically center the steering wheel which mimics the caster angle of the wheels in a real car (Wang, 2022). The gas and the brake pedals are converted to controls by taking the percentage of the maximum depression angle and using that as the percentage of throttle or brake in the simulation (Figure 13). All results from the brake were halved in order to make it harder for the operator to slow down and force a steering reaction (see section 3.4 for more).

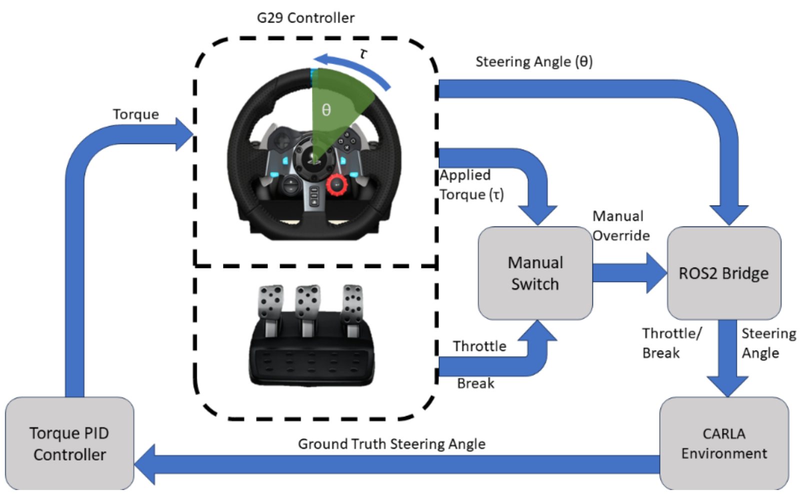


Figure 13: Control loop for the Logitech G29 steering wheel and how it feeds back to the simulator environment.

3.3.3 *Takeover*. Takeover was allowed at all times. However, we asked the participants to only take over when an emergency was imminent. This ensured they were using the ADAS and not unnecessarily taking over.

In order to make takeovers as smooth as possible, we needed the steering wheel to move as it does for the autonomous controller. This is accomplished by using same methods as auto-centering. We take the ground-truth steering information from CARLA and map it back to the steering wheel's force feedback via the driver (Dosovitskiy et al., 2017; Atsushi, 2023).

The operator has two ways to take over for the ADAS: first, depressing the brake more than ten percent of its full range; second, turning the steering wheel so force feedback is applying torque for 0.1 seconds. Once either condition is met, we hand full control to the operator and allow them to continue as they wish.

On top of handling simulation, CARLA allows us to create new objects like cameras. We create two new cameras, an RGB camera and a semantic segmentation camera. Both cameras are placed in the driver's seat of the car (Figure 14). Using the bridge and Open Computer Vision (OpenCV) we display these images to the user with three added pieces of information, an "A" or an "M" to let them know the control mode (Figure 15), the ready indicator (an "X" when the trial has not yet begun and an "O" once it has) (Figure 16), and the labels above the pedestrian models (Carla Simulator, 2022; Bradski, 2000).

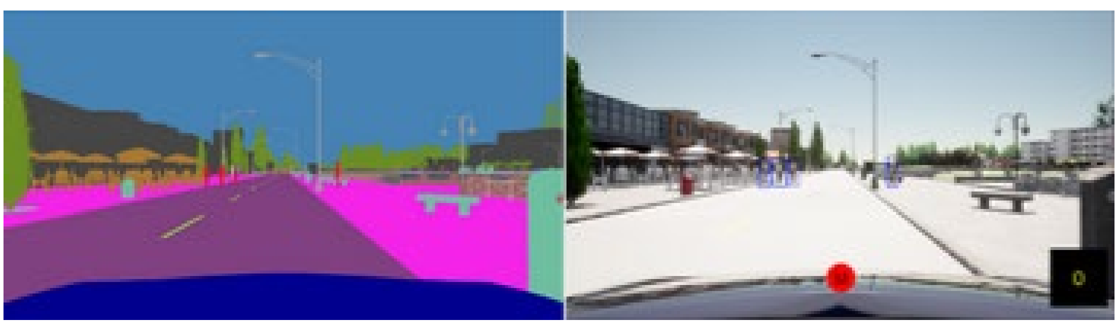


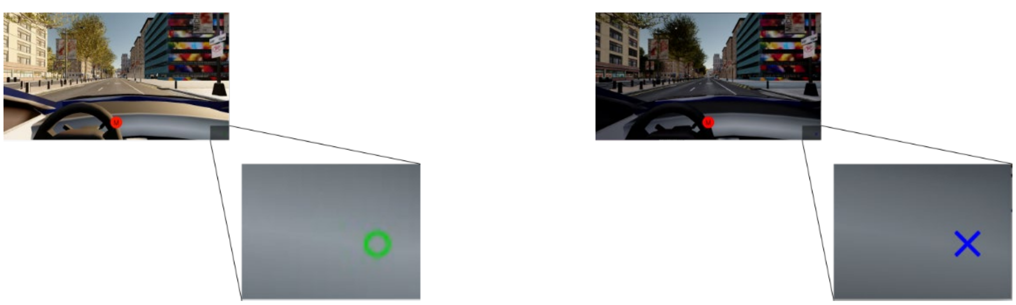
Figure 14. Comparison of the RGB and semantic segmentation camera.

3.3.4 User Interface. In the previously referenced Figure 12a the buttons labelled "A" are used for turning on autonomous. They were chosen due to their size and proximity to the steering wheel. The participant's hands did not need to leave the operating location in order to press the buttons.

Once the user sees the green "O" indicating they are ready to begin, they can press either of the paddle shifters to toggle on autonomous. The vehicle will accelerate to 15m/s (more on this choice in Section 3.4) along a predefined route. Once the vehicle is

35m away from the pedestrians, their label will appear and they begin to cross the street at 1.7m/s. It is at this point that the driver must react and make a decision. Either moving the steering wheel or depressing the brake will allow the participant to resume control and maneuver how they wish. Once they stop moving by passing the pedestrians' location, colliding with an obstacle, or other means, the trial is complete and all data and a video recording of the trial are saved. For a visual description of the interaction, see Figure 17.

Figure 15. RGB camera view of the scenario placed in the driver's seat.
In the lower center is the operational status. In the lower right corner is the "ready" indicator.



(a) Blue "X": The symbol that the scenario has not yet begun.

(b) Green "O": The scenario is ready to begin.

Figure 16. Symbol for stopping and beginning a trial.

3.3.5 Choosing a Location. As mentioned above, we noted that participants would remain in the same lane in which the trial began. Both scenarios where $\Delta_v = 0$ and similar scenarios in different locations help us to better account for this bias (Figure 18). The map used is a small city with several 4 lane avenues. Four lanes make it easy to run trials in both the right and left lane in an attempt to mitigate the preference.

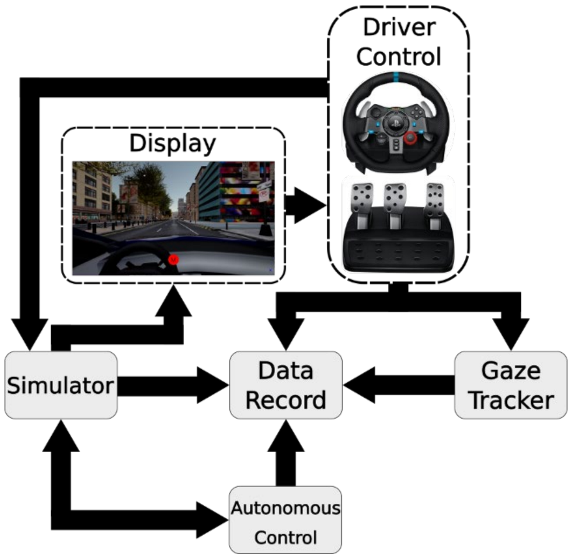


Figure 17. Flow diagram for how different parts of the simulation feed the recording.
The recording collects data from the user from the steering controller and the gaze tracker. From the simulator and the autonomous controllers, it collects information about the car itself (e.g., angular velocity, pedestrian position, etc.). The simulator also passes the user's view to the recorder which records each frame, so that we can watch the footage later.

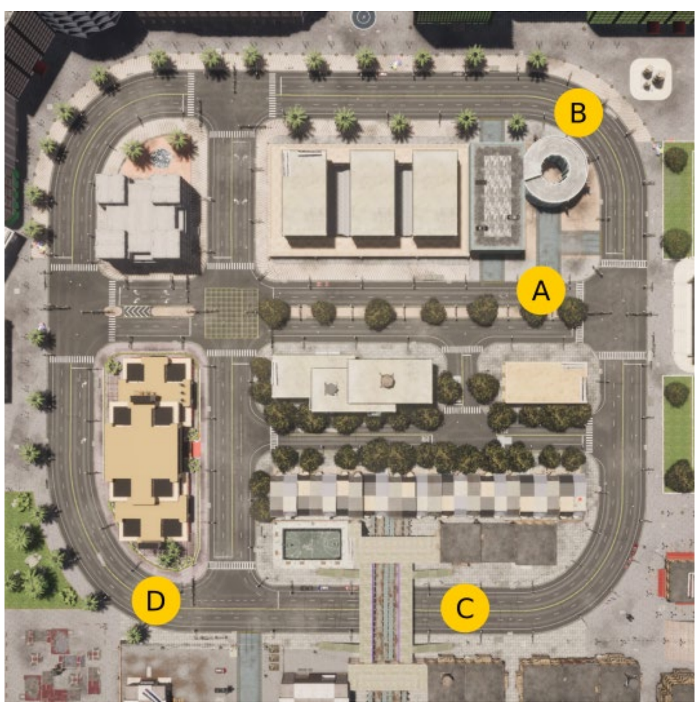


Figure 18. Bird's-eye view of the map used.
“A” and “B” are the two test locations. “C” and “D” are the training locations

When deciding how to design trials for a single participant, we began with a simple, straight street (Figure 19). As explained in the above paragraph, we also mirrored all trials from the right into the left lane, so we always have an even number of trials. This results in six trials: three from the different values of Δ_v and another three from the other side of the street. Since we could easily run twice the number of tests and keep the entire session to less than an hour, we decided to run all of our tests around a left-hand curve in the street (labelled "B" in previously referenced Figure 18). While straight streets are more similar to the questionnaires used in Dr. Cohen's previous PVT work, curved streets were an interesting way to see if people reacted in the same way when they could not see the destination. For these reasons, we decided to duplicate all trials on the left-hand curve to see if there is any difference in the reaction of participants (Figure 20).

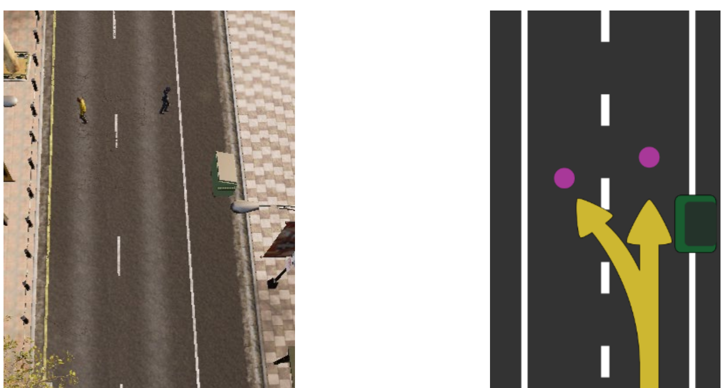
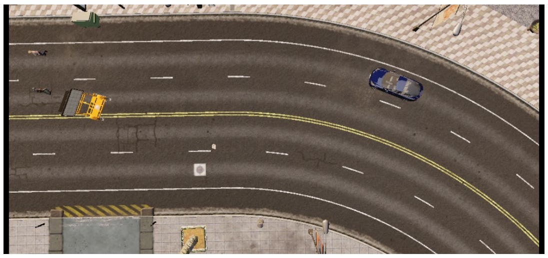


Figure 19. Top view of the straight trial course.

The parameters that are most difficult to translate to our scenarios are those for the car itself. In most trolley problems, the speed of the tram is only relevant to the extent that it causes the death of the unfortunate sacrifice. In our scenario, we needed to tune the parameters, so that both the driver has enough time to make a conscious choice and the speed is still high enough to cause harm to the pedestrian.



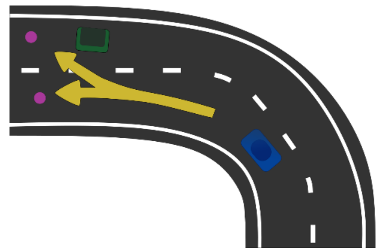


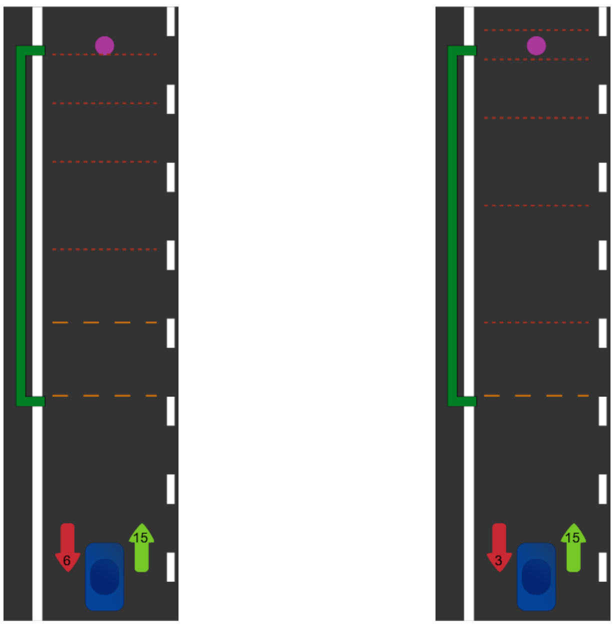
Figure 20. Top view of the curved trial course.

3.4 Assigning Parameters to the Vehicle.

The first parameter we considered was reaction time. The standard time for safely stopping before hitting an obstacle is known as the Safe-Stopping Time (SST). The standard for SST in the United States is 2.5s, although some take the position that this time should be longer (AASHTO, 2011; Wood and Zhang, 2021). We anticipated that our experiments would show SSTs less than 2.5 seconds because we told our participants to expect emergencies, so they were vigilant. In our pilot testing we saw similar reaction times, so we built our scenarios for those times.

Once we decided on a SST, the next parameter was vehicle speed. The CARLA map we were using has recommended speed limits for all of the streets. On the street in question the speed limit is 12.0m/s. We decided to use this as our default speed since it is consistent with the map as well as with the 25 mph speed limit found in small American cities (Board, 1998). Due to hardware issues, this speed was increased to 15m/s to compensate for rendering lag. This is discussed more in Chapter 5.

The Safe Stopping Distance (SSD) is 30m, which is the SST (2.5s) multiplied by the velocity (12m/s). We wanted to give the driver ample time to react so they could choose to which side to steer, but that also meant they had ample time to brake. Much of our interest is in whom the participant chooses to hit. If the participant is able to fully brake, thus hitting nothing, that is the clearly best choice. Because we do not want them to be able to fully brake in our experiment, we chose to weaken the brake by a factor of 2. Figure 21 shows the difference between these two braking strategies.



(a) A brake with normal strength. (b) A brake with half strength.

Figure 21. Green vertical bar is the triggering distance.
Each Horizontal represents a half second. The long, dashed bars represent time before the driver has reacted. The small dashes represent the time after the driver has depressed the brake.

3.5 Test Procedure

In accordance with IRB protocol (H23-0662), when running a participant through the experiment the process is done in 3 steps.

1. Intake
2. Training
3. Testing

3.5.1 Taking in a Participant. When a participant arrives, they are given a short presentation on what they will be completing. This includes the experiment's goals, what they will be doing, how to approach their tasks, and completing any necessary paperwork like a demographic questionnaire and the IRB consent form. Before beginning participants were told the following:

> We are looking to see how humans react in emergency scenarios with autonomous vehicles. [These reactions] will help us to model reactions and create shared control strategies that can improve outcomes for both pedestrians and the operator.

Our participants knew they would be taking part in scenarios where emergencies would occur. They were then given the demographic and consent forms and had the option to opt out. Described to the participants were also the format of the training scenarios and how to interpret the labels during the test scenarios. Lastly, they were given general information about wearing noise cancelling headphones, having their view of the rest of the room blocked by a screen to reduce distractions, ensuring proper calibration of the gaze tracker between each scenario, and how to decipher the begin and stop symbols (These symbols can be seen in previously referenced Figure 16 above).

3.5.2 Training. After the presentation was complete, participants were seated in a racing chair and given the ability to adjust it to their comfort before the training began (Figure 22). After adjusting, the initial gaze tracker calibration was performed.

The training course is a simple obstacle course repeated in four locations all similar to the locations in the test scenario referenced in Section 3.3.5. The training locations are labelled "C" and "D" in previously referenced Figure 18. The first four times a participant ran through these scenarios they were driving manually to get used to the simulator and the controller. Participants were instructed specifically to get a feel for the brake because it is weaker than in a normal vehicle. After completing all 4 scenarios, participants were given a chance to retry any of them until they self-reported feeling comfortable with manually operating the vehicle (Figure 23 and Figure 24).

Figure 22. OpenWheeler GEN 3 Racing Chair.

The four scenarios were then repeated, beginning with autonomous. The participants were instructed to focus on taking over for the autonomous controller when they had to avoid a crash. In these scenarios, they were told to focus on using the brake and steering wheel and feel comfortable with the effort required to take over. This process took about 10 minutes and was repeated until participants felt confident manually taking over for the ADAS. Limitations of this approach are discussed further in Chapter 5.

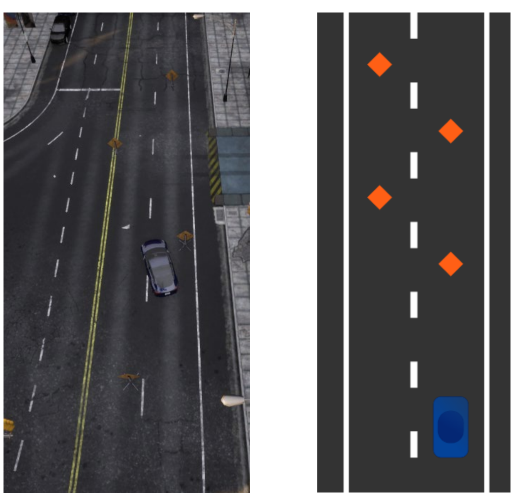


Figure 23. Top view of the straight training course.

Figure 24. Top view of the curved training course.

3.5.3 Testing. After training, participants moved directly into the test scenarios. As previously stated, there were 12 different scenarios given to each participant. In between each scenario, the gaze tracker's calibration was confirmed. The gaze tracker was re-calibrated if not within the range given by the gaze tracker's manufacturer (Gazepoint, 2021). Individual scenarios were organized into a Latin Square to reduce carry-over effect (Grant, 1948). The description of how these were defined can be found in Figure 25 and the Latin Square that was used can be found in Table 3. Since we had 36 participants we were able to run through the entire Table 3 times. Each trial was run as described above in Section 3.3.

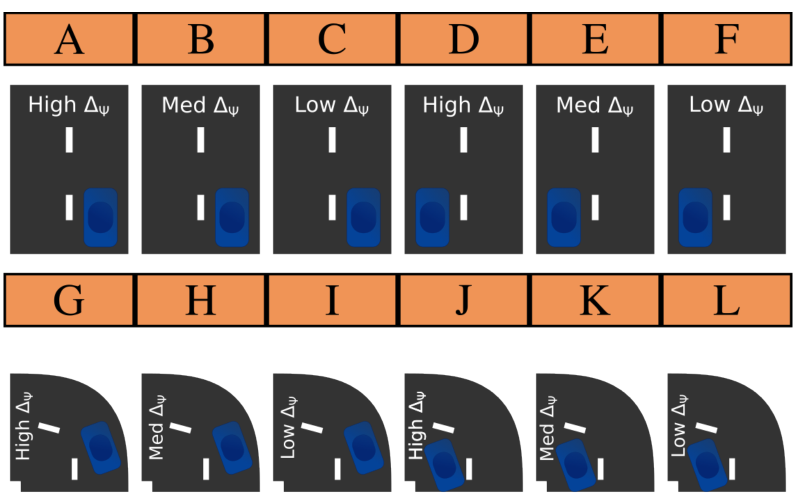


Figure 25. Labelling of each the Latin Square scenarios.

Table 3. The Latin Square Table Used.

A	B	L	C	K	D	J	E	I	F	H	G
B	C	A	D	L	E	K	F	J	G	I	H
C	D	B	E	A	F	L	G	K	H	J	I
D	E	C	F	B	G	A	H	L	I	K	J
E	F	D	G	C	H	B	I	A	J	L	K
F	G	E	H	D	I	C	J	B	K	A	L
G	H	F	I	E	J	D	K	C	L	B	A
H	I	G	J	F	K	E	L	D	A	C	B
I	J	H	K	G	L	F	A	E	B	D	C
J	K	I	L	H	A	G	B	F	C	E	D
K	L	J	A	I	B	H	C	G	D	F	E
L	A	K	B	J	C	I	D	H	E	G	F

3.6 Data Collection

In these experiments we collected many data. From the participant we collected:

- The x,y coordinate of their gaze on the screen (Figure 26)
- The angle, angular, and velocity of the steering wheel
- How much the brake and throttle were depressed

We also were able to obtain ground-truth data from the simulator (Figure 27). These include:

- The vehicle position in the world frame
- The vehicle's velocity (angular and linear)
- If the vehicle was being controlled by the ADAS
- Camera information like the focal length and field of view
- The pedestrians' position in the world frame

- The pedestrians' projected position onto the driver's view
- The label and model used for each pedestrian

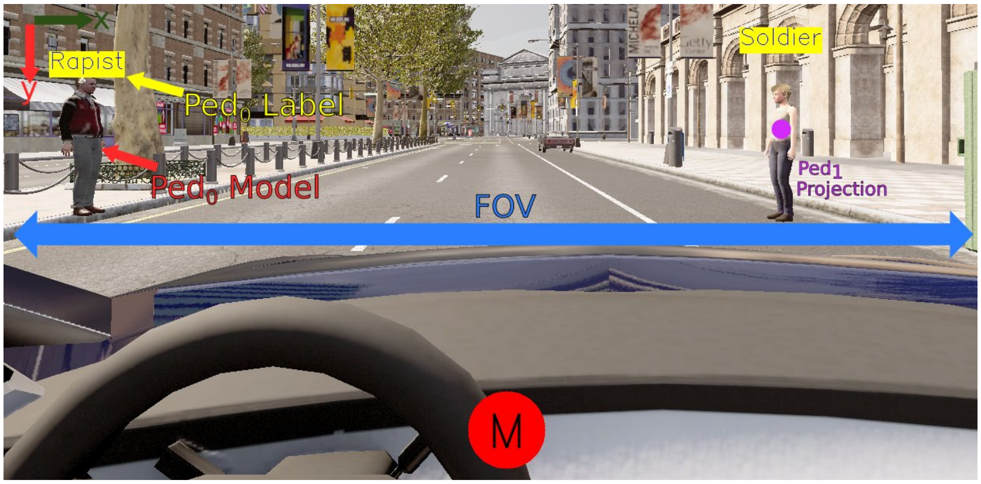


Figure 26. Data collected from the driver view.

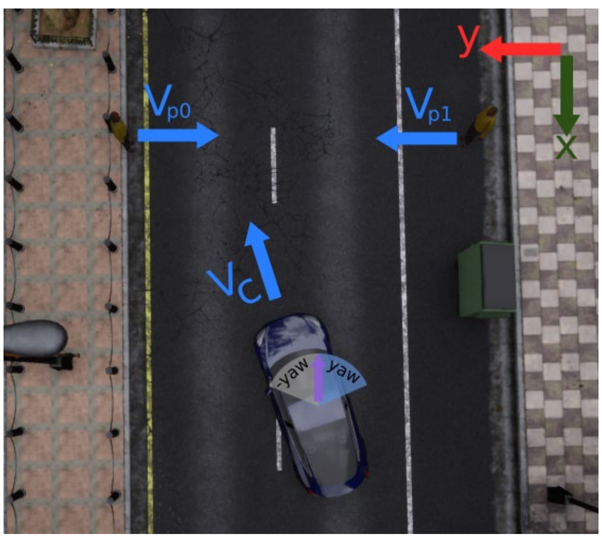


Figure 27: Data collected from the simulation.

From these data we also hope to be able to model and predict the driver's reaction using system identification. This requires the use of gaze data and the steer reaction.

CHAPTER 4: RESULTS

4.1 Summary and Outliers

In total we had 36 participants, we collected ages in ranges. Six participants were in the 18-21 category, seven in the 22-24 ranges, three were in the 25-28 range, one was between 29 and 32, and nineteen were above 32. 17 of our participants were female and 19 male. We only had 3 participants report as left-hand dominant and 1 participant that reported as ambidextrous. Each participant ran through 12 trials, but we lost the data for one of the trials. In total this left us with 431 trials. An example of how the trials look as controlled by the participant can be seen in Figure 28.

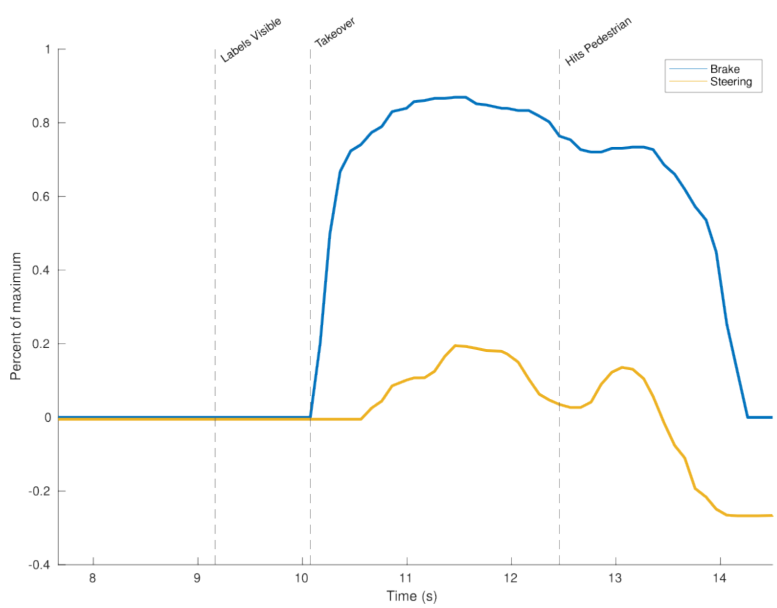


Figure 28. Example of the control inputs of one trial from one participant.
Each of the vertical grey lines represents an important time in the recording. The difference between the first two bars is the reaction time. The distance between the second and third lines is where most of the data is collected like gaze and steering. The last vertical line is the point at which the driver has solidified their choice and where the yaw is recorded.

There was a bias in scenario generation to higher value targets on left side of the trial (positive Δ_v). Of the total 431, we would expect 144 of the randomly generated scenarios to be positive, negative, or zero. Instead of the 431 only 102 scenarios had a negative relative value (Figure 29).

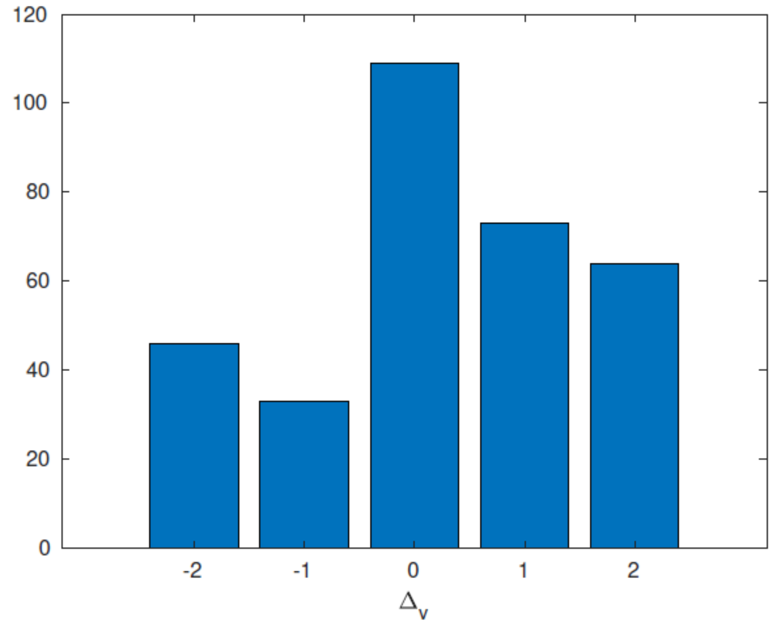


Figure 29. Number of scenarios run with each relative value. Note that $\Delta_v = 1$ is higher than $\Delta_v = -1$.

Because the vehicle was always going to hit a pedestrian, participants quickly learned that it was possible to avoid hitting anyone by pressing the brake early. Participants were instructed to avoid reacting until they saw the labels as an attempt to prevent this from occurring. Even without predictive braking, reaction times were faster than the standard in AASHTO (2011). This was not a surprise as drivers were alert and expected to have to brake quickly. Čulík et al. (2022) found that in these cases in a simulator the lower bound for reaction time was 0.4s. So, we ignore all data that have reaction time less than 0.4s. In total this leaves us with 366 trials. In Figure 30, we see all the 366 trials and what the driver made.

4.2 Important Values

We identified nine distinct features that we wanted to further analyze. The correlation matrix for each of these features is shown in Figure 31.

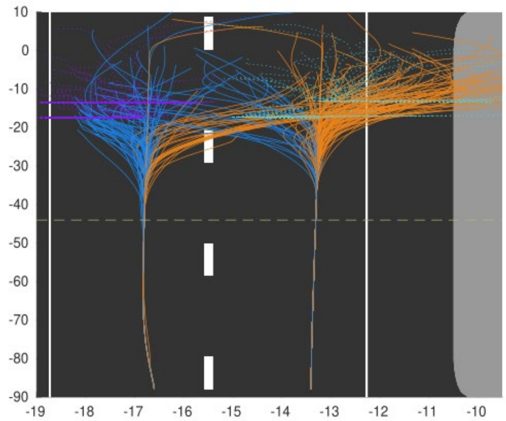
1. Start Lane — The lane in which the participant began the trial (-1 is left, 1 is right).
2. Street Type — Whether the participant began on a straight (-1) or curved (1)

street.

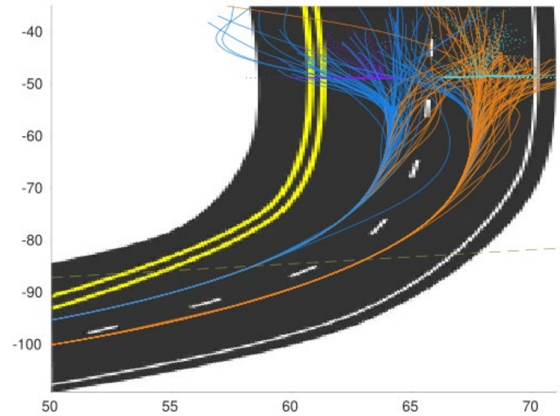
3. Δ_v — The difference of value of the trial. A positive number means the pedestrian on the left has a higher value. Larger absolute values of Δν mean smaller overlaps between the values of the characters in the scenarios.
4. Reaction Time—The time between when the pedestrian starts moving and when the participants take over for autonomous.
5. Choices — The pedestrian the participant chose to hit (-1 for left, 1 for right). Note participants may not have turned the wheel at all, continuing straight is considered a choice for the lane they are currently in.
6. Maximum Right Turn Action— The largest angle the steering wheel turned to the right (this will always be a positive number)
7. Maximum Left Turn Action — The largest angle the steering wheel turned to the left (this will always be a negative number)
8. Yaw Angle — At the time the vehicle passes the pedestrian, what direction is it facing (in degrees 0 is straight, -90 is 90o to the left, 90 is 90o to the Right).
9. Gaze — The average difference of the distance of the X component (horizontal motion) of gaze from the bounding box surrounding each pedestrian (negative to the left, positive to the right) (Figure 32).

4.3 Bias

The high correlation between the start lane and choice reveals a bias to stay in the current lane rather than make a choice (Figure 33a). Part of this is likely because for the $\Delta_v = 0$ cases the most reasonable choice is to continue in your current lane as PVT does not offer a *correct* choice. This only accounts for 1/3 of the decisions, looking at the choice based on Δ_v (Figure 33b) also shows that drivers would rather stay in their lane regardless of the value.



(a) An image of all trials on the straight street.



(b) An image of all trials on the curved street.

Figure 30: Trials where the driver is directed.
The blue lines direct toward the left pedestrian, orange to the right. The purple and light teal lines are the pedestrians. The dotted line in the center is when the labels on the pedestrians become visible and we begin to measure reaction. Reaction times below 0.4s have been removed.

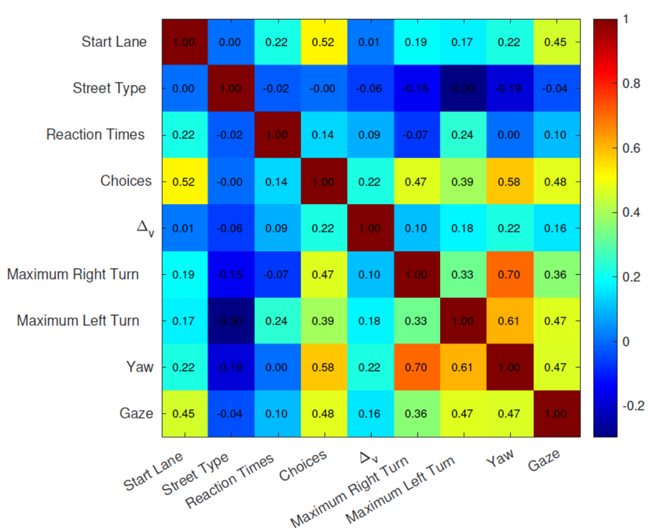


Figure 31. Correlation between the nine features.

Start lane and gaze are correlated in a similar way to the choices. Drivers are more likely to spend time looking straight ahead when continuing along a road (Fan et al., 2021). They are, therefore, more likely to spend time looking at the pedestrian directly in front of them.

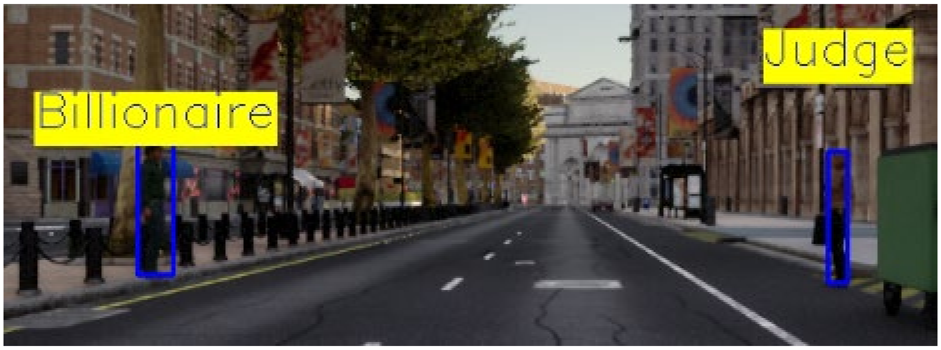
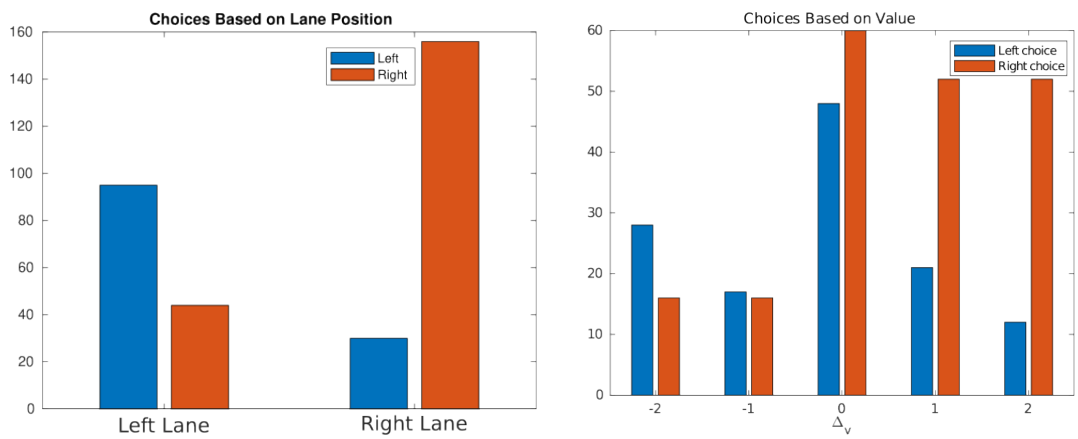


Figure 32. Image showing the bounding boxes surrounding the pedestrians.



(a) A bar graph showing the relationship between choice and starting lane. Note the bias to choose to stay in one's own lane.

(b) A bar graph showing the relationship between choice and Δ_v. Note the tendency to choose to hit the pedestrian with a lower value, as predicted by PVT.

Figure 33. Bar graphs showing the relationship between choice and other categories.

The street type also presents a bias specifically in the steering data. The turn for the trials was always to the left, so larger left turns are expected.

4.4 Reaction Time

PVT predicts that with a higher overlap of value the reaction time should increase. We would then have expected a positive correlation between $|\Delta_v|$ and choice, but there is none. To show this a one-way ANOVA was conducted to determine the effect of $|\Delta_v|$ on reaction time. The results indicate no effect, [$F(2, 322) = 0.07$, $p=0.935$] (Figure 34). We, therefore, fail to reject the null hypothesis that $|\Delta v|$ has an effect on reaction time. We expect that regardless of the scenario the participant was in, as soon as the label was visible, they pressed on the brake or began to steer. After doing so they decided what to do. Figure 35 shows an example of braking early versus steering early. In the red line, the participant immediately steers to the center of the road before deciding to collide with the right most pedestrian. The green shows no direction change and instead braking before deciding what pedestrian to hit.

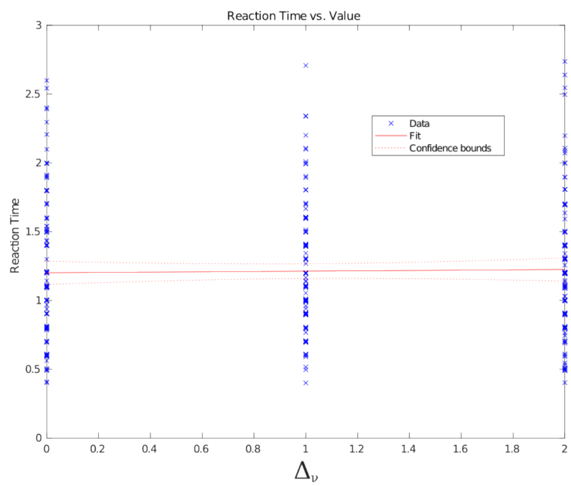


Figure 34: Relationship between reaction time and $|\Delta_v|$.
Contrary to our hypothesis, note the lack of any correlation between $|\Delta_v|$ and reaction time

Reaction time in these cases then, is different than the reaction time described in Cohen et al. (2022). The point of reaction in those experiments is when the participant makes a choice. In our work we are only measuring the point when the operator first reacts. As shown above a choice is made after the initial reaction.

4.5 Choices

The main idea behind PVT is that we value different people different amounts in their relationship to each other. Previously, this was shown through static questionnaires. These results show this holds even in these more complex, dynamic scenarios. Despite the bias toward remaining in one lane and the additional right-turn bias. The correlation between the two is not insignificant. A one-way ANOVA was conducted to determine the effect of Δ_v on choice. The results indicate a significant effect, [$F(4, 320) = 6.03$, $p = .0001$]. Post Hoc tests were conducted using Tukey's Honest Significant Difference

(HSD) procedure. The comparison revealed significant difference between $\Delta_v = -2$ (M=-0.1739; SD=0.9956) and $\Delta_v = 1$ (M=0.3973; SD=0.9241) (P-value = 0.011469, C.I. = [-1.056, -0.086342]); between $\Delta_v = -2$ and $\Delta_v = 2$ (M=0.6250; SD=0.7868) (P-value = 0.00012, C.I. = [-1.2967, -0.30108]); between $\Delta_v = 0$ and $\Delta_v = 2$ (M=0.1193; SD=0.9974) (P-value = 0.0060, C.I. = [-0.91131, -0.10015]) (Figure 36). We, therefore, reject the null hypothesis that different differences in value have the same effect on which way the driver will steer after taking over.

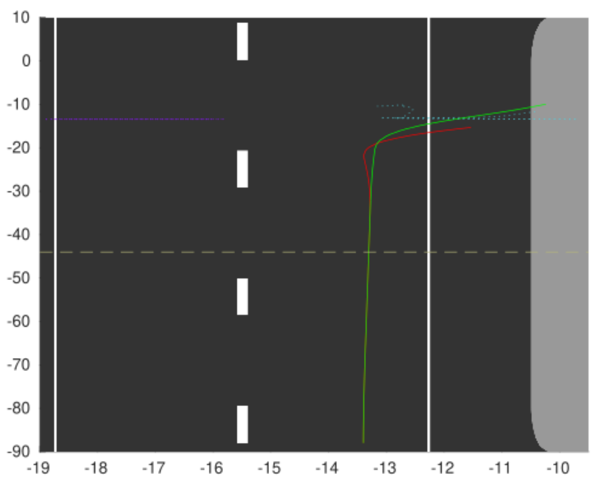


Figure 35. Two scenarios in which the participant reacts early but doesn't make a choice until much later. In the red example the participant steers to the middle of the lane before choosing. In the right, the participant brakes before choosing

Gaze, as we discussed earlier Pugeault and Bowden (2015), showed that they were able to predict the driver's decisions based on their gaze patterns. In order to determine significance, values of gaze greater than 1 (off the screen) were removed. A one-way ANOVA was conducted to determine the effect of choice on gaze. The results indicate a strong relationship [$F(1, 319) = 106.07$, $p < 0.0001$]. Post Hoc tests were conducted using Tukey's Honest Significant Difference (HSD) procedure. The comparison revealed significant difference between choosing to hit the left pedestrian (M=-0.0700; SD=0.1257) and choosing to hit the right pedestrian (M=0.0710;

SD=0.1147) (P-value< 0.0005, C.I. = [-0.16792, -0.11422]) (Figure 37a). We, therefore, reject the null hypothesis that different differences in value have the same effect on at which pedestrian the participant spends more time looking. Similarly, a one-way ANOVA was conducted to determine the effect of Δ_v on gaze. The results indicate an effect, though not quite significant [F(4, 316) = 2.17, p = .0726]. PVT expects that cases where $\Delta v = 0$ to be random this will make the effect of Δ_v on gaze seem lesser. Removing values where $\Delta v = 0$ reveals a stronger relationship [F(3, 208) = 2.97, p = .0327]. Post Hoc tests were conducted using Tukey's Honest Significant Difference (HSD) procedure. The comparison revealed significant difference between $\Delta_v = -2$ (M=-0.0258; SD=0.1623) and $\Delta_v = 2$ (M=0.0455; SD=0.1346) (P-value = 0.0350, C.I. = [-0.1391, -0.0035]) (Figure 37b). We, therefore, reject the null hypothesis that different differences in value have the same effect on which pedestrian the participant spends more time looking at.

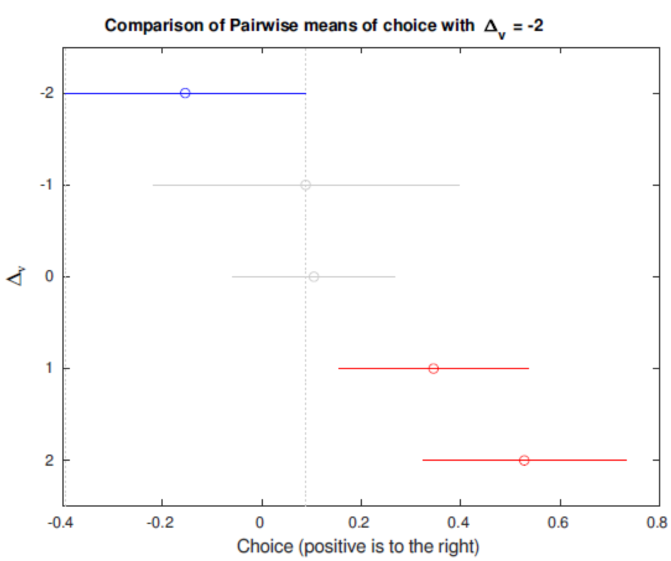
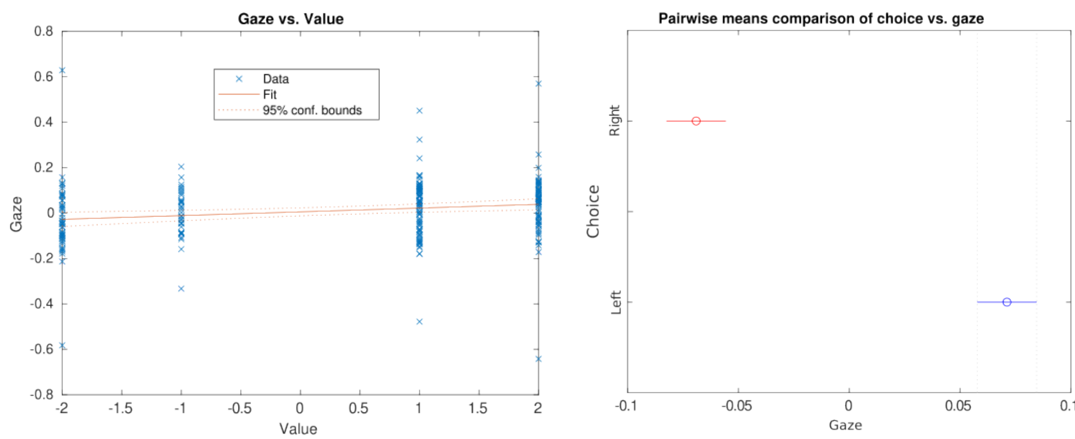


Figure 36. Choice against the Δ_v of the scenario.
Again, this relationship is what we would expect from PVT. The relationship between value and choice holds; participants more often chose to hit the lower value pedestrian and did so more often when the value difference was higher.

4.6 Gaze

These results are indicative of a compensatory task. As the operator identifies and gazes toward the lower value target, they turn the vehicle toward it centering it in their view. They remain looking at it while continuing and ensure they will hit the lower value pedestrian rather than avoiding the pedestrian with higher value.



(a) Choice against the average difference of the participants gaze with the center of the pedestrians' bounding box.

(b) The relationship of gaze vs Δ_v. Shows a positive correlation, those much of this can be explained with the stronger relationship between gaze and choice.

Figure 37. Relationship between Gaze and the two categories with which it is more closely related.

4.7 Yaw

The yaw is the direction the car is facing, this is measured at the time of collision. For most cars, the yaw is strongly correlated with the choice. To show this a one-way ANOVA was conducted to determine the effect of choice on the vehicle's yaw at the time of collision. The results indicate a significant effect, $[F(1, 323) = 188.11, p< 0.0001]$. Post Hoc tests were conducted using Tukey's Honest Significant Difference (HSD) procedure. The comparison revealed significant difference between choosing to hit the left pedestrian (M=-7.4977; SD=10.0242) and choosing to hit the right pedestrian (M=8.3709; SD=10.1978) (P-value< 0.0005, C.I. = [-18.041, -13.518]) (Figure 38). We, therefore, reject the null hypothesis that drivers will be heading in the same direction at the time of collisions regardless of which pedestrian they chose to hit. These results indicate that operators do not begin to correct the vehicles steering until they have already collided with the pedestrian, they intended to hit rather than deciding who they would hit and correcting back to avoid running off the road. This suggests that their decision is more immediate and that they want to ensure their choice is completed before addressing what will happen next.

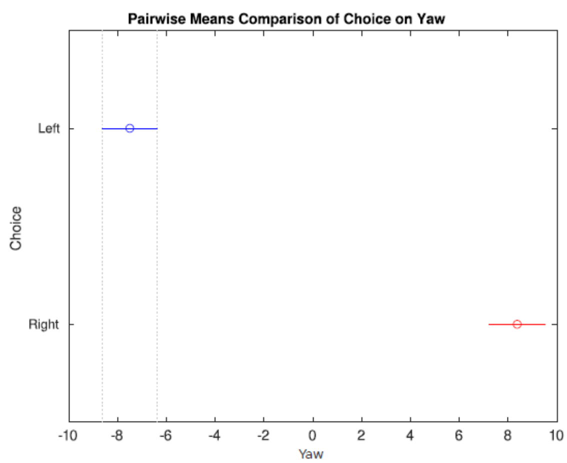


Figure 38. Pariwise comparison of choices vs. Yaw.

CHAPTER 5: DISCUSSION

5.1 Limitations

The first limitation is attributable to the simulation. Our participants knew that no harm would come to them or the pedestrians, so their actions may not have reflected how they would act in a real-world environment. Despite being told to act as if they were driving a real car, several participants reported feeling as though they were in a video game.

When creating the scenario, we experienced frame rate issues. The scenario was running at ≈16 fps instead of the intended 20 fps. For this reason, as mentioned earlier, we increased the speed of the ACC to 15 m/s from the original 12 m/s. This gave participants the same time to react as if they were operating at 12 m/s at full speed instead of a simulation running 1.2 times slower.

Our participant pool covered a wide range of ages. Because of this and the frame rate issues, some of the older participants reported struggling to read the labels – especially on the first trial when they were not expecting to see them. Particularly in their early trials, we found that a few of the participants remained in their own lane due to an inability to distinguish between pedestrians.

Hardware was again a limiting factor when running scenarios because we only had one monitor. Participants were not able to look around like they would in a normal car, so their field of view was lessened. Further, they were not given any mirrors or sound to gather information about their surroundings. Though these were not necessary for the trials, some participants reported missing the additional sensory information.

As discussed in Chapter 4, scenario generation was biased. More $\Delta_v = 1$ scenarios were generated than $\Delta_v = -1$. Partly because of this bias, we saw a tendency for the drivers to hit the pedestrian in the right lane. The right was further biased out of self-preservation. The left side of the road has a wall preventing any further turns while the right only has a large dumpster.

Lastly, the environment itself led to several early reactions and unrepresentative behavior. Many participants used the brake before being presented with an emergency. They understood after a few trials they were often going to hit the pedestrians and were doing their best to avoid this. In our best effort to minimize these early reactions, participants were instructed to allow the autonomous to drive until they had to take over.

5.2 Conclusion

In this study we measured drivers' reactions to sacrificial dilemmas and attempted to use PVT to explain the choices made. Not all results conformed with the existing model. In the work of Cohen et al. (2022), reaction time has a strong relationship with $|\Delta_v|$, but in our work this is not the case. In our experiment, the time when you make a choice of whom to sacrifice is different from the time you first react. However, in the static PVT experiments these are the same. The natural reaction when in a car is to brake. This helps to give the operator more time to react as well as the best chance of avoiding any collision no matter who is present. When drivers did make a choice to collide with a pedestrian, they did follow our expectations, hitting the one with the lowest value. Further, they did this more consistently when the difference between the two values was largest.

Variables not explained by PVT also had strong power in outcomes, corroborating with Pugeault and Bowden (2015) drivers tended to look where they were going. In our case, that is at the pedestrian they were going to hit. Going forward, we hope to be able to use this data to predict the drivers' goals and how they react and determine value in dynamic, time sensitive scenarios.

CHAPTER 6: FUTURE WORK

We generated much data in this project and have not yet had time to process it all. We will, in the coming months, be looking into modeling the actions of our participants. We will use different system identification techniques to better understand how operators react in emergency scenarios. We have begun some promising preliminary work on gaze frequency and absolute value but are yet to be able to fully explain the results (Figure 39).

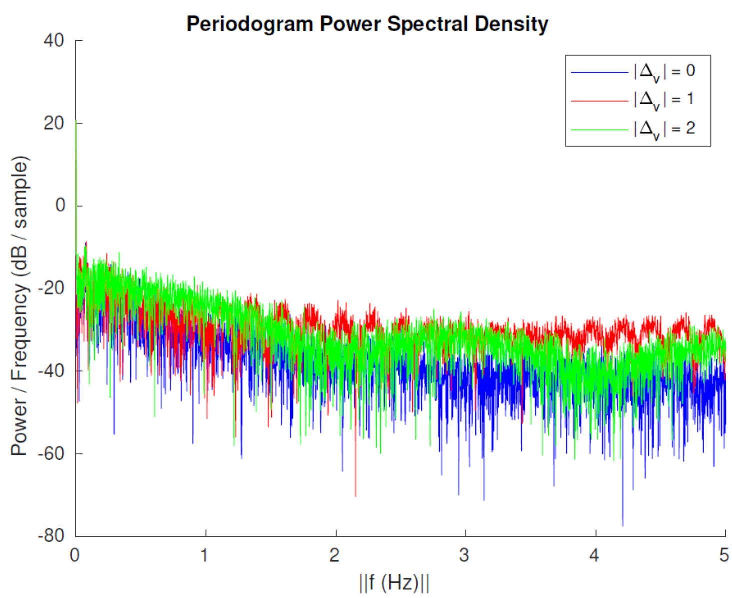


Figure 39. Graph of the Power Spectral Density (PSD) against the frequency. Data obtained using fast Fourier transform of gaze separated by the absolute difference in psychological value of the scenario.

In order to better measure the reaction of our participants in silent failures, one of our hopes during the semester was to record a suite of videos where the ADAS randomly succeeds or fails. These videos wouldn't have a brake option and instead would only allow the user to choose to stay in a lane or turn into the other. These videos could be used on a larger group of participants since they could be done at home and would allow us to test whether reaction time is dependent on the pedestrians being present in a more controlled scenario.

The above can be extended to include success and failures in a completely dynamic scenario similar to the one we currently have. This would give us a benefit similar to delaying reaction time as the emergency will be a surprise. This will also give us insight into whether the reactions change when the takeover is more unexpected. We did our best to control for it in these experiments, but a surprise takeover might yield different results. Along with this, we can change from silent to loud failures. We could then see how the changes impact participants' reactions and if it improves outcomes.

There are other adjustments that can be made easily to test the value judgments made by people when they lack time to react. For example, the findings in Bonnefon et al. (2016) are not in perfect alignment with those in Cohen et al. (2022). In the former, participants reported rather running over a single person in general, whereas in the latter, group size did not have much effect. The questions posed in the two questionnaires are different, Bonnefon et al. (2016) asked if the passenger of an SDC should be sacrificed when more than one person could be saved, where Cohen et al. (2022) asked whether you would choose to kill one or a group of the other. We have the ability to change the behavior of a car to make these choices and see if our participants override the action of the vehicle based on value or amount.

Another promising area of study is to use a more immersive simulation. There are other examples of CARLA being used in virtual reality (Figure 40), thus giving our participants a wider field of view and making the environment more realistic. We can accomplish this by building off the work of Silvera et al. (2022) who have been able to use their virtual reality headsets to better capture gaze and build a more immersive environment than a single computer monitor can offer.

Figure 40. Physical setup with participant driver alongside experimenter's monitoring setup (Silvera et al., 2022).

REFERENCES

AASHTO (2011). *A Policy on Geometric Design of Highways and Streets, 2011*. AASHTO. Google-Books-ID: puLERKfS0RcC.

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.

Atsushi, K. (2023). Ros g29 force feedback.

Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., Jesus, L., Berriel, R., Paix˜ao, T. M., Mutz, F., de Paula Veronese, L., Oliveira-Santos, T., and De Souza, A. F. (2021). Self-driving cars: A survey. *Expert Systems with Applications*, 165:113816.

Baikejuli, M., Shi, J., and Qian, Q. (2023). Mobile phone use among truck drivers: The application and extension of the theory of planned behavior. *Accident Analysis & Prevention*, 179:106894.

Banks, V. A., Eriksson, A., O'Donoghue, J., and Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, 68:138–145.

Bartels, D. M. and Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121(1):154–161.

Beringer, D. B., Harris, H. C., and United States. Office of Aerospace Medicine (1997). Automation in General Aviation: Two Studies of Pilot Responses to Autopilot Malfunctions. Technical Report DOT/FAA/AM-97/24, FAA.

Bianchi Piccinini, G., Lehtonen, E., Forcolin, F., Engström, J., Albers, D., Markkula, G., Lodin, J., and Sandin, J. (2020). How do drivers respond to silent automation failures? Driving simulator study and comparison of computational driver braking models. *Human Factors*, 62(7):1212–1229.

Board, T. R. (1998*). Managing Speed: Review of Current Practices for Setting and Enforcing Speed Limits – Special Report 254*. The National Academies Press, Washington, DC.

Bonnefon, J.-F., Shariff, A., and Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science, 352*(6293):1573–1576. Publisher: American Association for the Advancement of Science.

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Brandenburg, S. (2021). Human Factors of Automated Driving Systems. In Lentin, J. and Mondal, A. K., editors, *Autonomous Driving and Advanced Driver-Assistance Systems* (ADAS), pages 321–334. CRC Press, Chapter 14. Num Pages: 14.

Brown, B. and Laurier, E. (2017). The Trouble with Autopilots: Assisted and Autonomous Driving on the Social Road. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 416–429, New York, NY, USA. Association for Computing Machinery.

Carla Simulator (2022). Ros bridge.

Chakraborty, S., Laware, H., Castanon, D., and Zekavat, S. R. (2016). High precision localization for autonomous vehicles via multiple sensors, data fusion and novel wireless technologies. In *2016 IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference* (UEMCON), pages 1–9.

Christov-Moore, L., Conway, P., and Iacoboni, M. (2017). Deontological Dilemma Response Tendencies and Sensorimotor Representations of Harm to Others. *Frontiers in Integrative Neuroscience*, 11:34.

Cohen, D. J. and Ahn, M. (2016). A subjective utilitarian theory of moral judgment. *Journal of Experimental Psychology: General*, 145(10):1359–1381. Place: US Publisher: American Psychological Association.

Cohen, D. J., Cromley, A. R., Freda, K. E., and White, M. (2022). Psychological value theory: The psychological value of human lives and economic goods. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 48(12):2015–2050.

Counterpoint, T. (2022). ADAS Penetration Crosses 70% in US in H1 2022, Level 2 Share at 46.5%.

Cozzens, T. (2022). DeepRoute.ai completes L4 driverless test in busy Shenzhen, China.

Day, M., Norman, P., Poulter, D., O¨ zkan, O., and Rowe, R. (2023). The adoption and application of Intelligent Speed Assistance by private motorists: User and non-user perspectives. *Transportation Research Part F: Traffic Psychology and Behaviour*, 99:262–273.

Doll, S. (2023). GM looks to catch up in China after getting OK to begin testing autonomous driving tech.

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. (2017). CARLA: An Open Urban Driving Simulator. arXiv:1711.03938 [cs].

Drexler, D. A., Tak´acs, A., Galambos, P., Rudas, I. J., and Haidegger, T. (2018). Handover Process Models of Autonomous Cars Up to Level 3 *Autonomy*. In *2018*

IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI), pages 000307–000312. ISSN: 2471-9269.

Drexler, D. A., Tak´acs, A., Nagy, T. D., Galambos, P., Rudas, I. J., and Haidegger, T. (2019). Situation Awareness and System Trust Affecting Handover Processes in Self-Driving Cars up to Level 3 Autonomy. In *2019 IEEE International Work Conference on Bioinspired Intelligence* (IWOBI), pages 000179–000184.

Epic Games (2019). Unreal engine.

Epple, S., Roche, F., and Brandenburg, S. (2018). The Sooner the Better: Drivers' Reactions to Two-Step Take-Over Requests in Highly Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1):1883–1887. Publisher: SAGE Publications Inc.

Eriksson, A. and Stanton, N. A. (2017). Driving performance after self-regulated control transitions in highly automated vehicles. *Human Factors*, 59(8):1233–1248.

Fan, X., Wang, F., Song, D., Lu, Y., and Liu, J. (2021). Gazmon: Eye gazing enabled driving behavior monitoring and prediction. *IEEE Transactions on Mobile Computing*, 20(4):1420–1433.

Fox, E. (2021). Tesla Likely to Achieve Level 4 Autonomy in 2022, Says Elon Musk.

Fu, E., Hyde, D., Sibi, S., Johns, M., Fischer, M., and Sirkin, D. (2020). Assessing the effects of failure alerts on transitions of control from autonomous driving systems. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, page 1956–1963.

Gazepoint (2021). Gp3 eye-tracker.

Gill, T. (2020). Blame It on the Self-Driving Car: How Autonomous Vehicles Can Alter Consumer Morality. *Journal of Consumer Research*, 47(2):272–291.

Gold, C., Damb¨ock, D., Lorenz, L., and Bengler, K. (2013). "take over!" how long does it take to get the driver back into the loop*? Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1):1938–1942.

Goodall, N. J. (2014). Machine Ethics and Automated Vehicles. In Meyer, G. and Beiker, S., editors, *Road Vehicle Automation*, pages 93–102. Springer International Publishing, Cham.

Grant, D. A. (1948). The latin square principle in the design and analysis of psychological experiments. *Psychological Bulletin*, 45(5):427–442.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment.

Science (New York, N.Y.), 293(5537):2105–2108.

Hofbauer, M., Kuhn, C. B., P¨uttner, L., Petrovic, G., and Steinbach, E. (2020). Measuring Driver Situation Awareness Using Region-of-Interest Prediction and Eye Tracking. In *2020 IEEE International Symposium on Multimedia* (ISM), pages 91–95, Naples, Italy.

Jagacinski, R. and Flach, J. (2003a). Control Theory for Humans: Quantitative Approaches To Modeling Performance. In *Books Authored by Wright State Faculty/Staff*, page 65. CRC Press.

Jagacinski, R. and Flach, J. (2003b). Control Theory for Humans: Quantitative Approaches To Modeling Performance. In *Books Authored by Wright State Faculty/Staff*, page 105. CRC Press.

Jin, M., Lu, G., Chen, F., Shi, X., Tan, H., and Zhai, J. (2021). Modeling takeover behavior in level 3 automated driving via a structural equation model: Considering the mediating role of trust. *Accident; Analysis and Prevention*, 157:106156.

Lee, C., Seppelt, B., Reimer, B., Mehler, B., and Coughlin, J. F. (2019). Acceptance of Vehicle Automation: Effects of Demographic Traits, Technology Experience and Media Exposure. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1):2066–2070. Publisher: SAGE Publications Inc.

Lieu, H. and Koppa, R. J. (1999). Traffic-flow theory. Public Roades, 62(4).

Macenski, S., Foote, T., Gerkey, B., Lalancette, C., and Woodall, W. (2022). Robot operating system 2: Design, architecture, and uses in the wild. *Science Robotics*, 7(66):eabm6074.

Magazine, A. (2017). Audi AI traffic jam pilot.

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., and Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27:274–282.

Mole, C., Pekkanen, J., Sheppard, W., Louw, T., Romano, R., Merat, N., Markkula, G., and Wilkie, R. (2020). Predicting takeover response to silent automated vehicle failures. *PLOS ONE*, 15(11):e0242825.

Molla, R. (2021). Self-driving cars: The 21st-century trolley problem.

NHTSA (2016). Automated Driving Systems 2.0: A Vision for Safety.

NHTSA (2023). Early Estimate of Motor Vehicle Traffic Fatalities in 2022.

Pugeault, N. and Bowden, R. (2015). How much of driving is preattentive? *IEEE Transactions on Vehicular Technology*, 64(12):5424–5438.

SAE International (2021). J3016 202104: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International.

Samak, C. V., Samak, T. V., and Kandhasamy, S. (2020). Control strategies for autonomous vehicles. *CoRR*, abs/2011.08729.

Santos, G. H. G. and Larocca, A. P. C. (2019). Drivers Take-Over Performance from Partial Automation to Manual Driving. *Proceedings of the 2019 33th Congresso de Pesquisa e Ensino em Transporte da ANPET*, pages 3885–3888.

Shladover, S. E. (2016). The Truth About "Self-Driving" Cars. *Scientific American*, 314(6):52–57. Publisher: Scientific American, a division of Nature America, Inc.

Silvera, G., Biswas, A., and Admoni, H. (2022). Dreyevr: Democratizing virtual reality driving simulation for behavioural & interaction research. *CoRR*, abs/2201.01931.

Soares, S., Lobo, A., Ferreira, S., Cunha, L., and Couto, A. (2021). Takeover performance evaluation using driving simulation: a systematic review and meta-analysis. *European Transport Research Review*, 13(1):47.

Sugeno, M. and Murakami, K. (1984). Fuzzy parking control of model car. In *The 23rd IEEE Conference on Decision and Control*, pages 902–903.

SWOV (2012). The Relation Between Speed and Crashes. Fact Sheet, SWOV - Institute for Road Safety Research, Leidschendam, the Netherlands.

U.S. (2012). U.s. code title 14 - aeronautics and space.

Čulík, K., Kalašová, A., and Štefancová, V. (. (2022). Evaluation of driver's reaction time measured in driving simulator. *Sensors (Basel, Switzerland*), 22(9):3542.

Wang, Z. (2022). Methods in research about effect of caster angle on vehicle steering and stability. *Highlights in Science, Engineering and Technology*, 13:45–53.

Wood, J. S. and Zhang, S. (2021). Evaluating Relationships between Perception-Reaction Times, Emergency Deceleration Rates, and Crash Outcomes using Naturalistic Driving Data. *Transportation Research Record: Journal of the Transportation Research Board*, 2675(1):213–223.

Zanchin, B. C., Adamshuk, R., Santos, M. M., and Collazos, K. S. (2017). On the instrumentation and classification of autonomous cars. *In 2017 IEEE International Conference on Systems, Man, and Cybernetics* (SMC), pages 2631–2636.

Zhang, W., Zhao, Y., Zhang, X., and Lin, F. (2020). Adaptive shared control strategy for lane changing assistance system via multi-mode switching. *Journal of the Franklin Institute*, 357(18):13304–13325.

Zhang, Y., Zhou, X., Wang, W., Hu, Y., and Sun, X. (2023). Keeping in the lane! Investigating drivers' performance handling silent vs. alerted lateral control failures in monotonous partially automated driving. *International Journal of Industrial Ergonomics*, 95:103429. 44