Impact of Highly Automated Vehicle (L4/5 AV) External Communication on Other Road User Behaviors

October 2022

Final Report

SAFE-D
SAFETY THROUGH DISRUPTION

TRANSPORTATION INSTITUTE
VIRGINIA TECH.

San Diego State University
Leadership Starts Here

Texas A&M Transportation Institute

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. VTTI-00-027	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Impact of Highly Automated Vehicle (L4/5 AV) External Communication on Other Road User Behavior		5. Report Date October 2022 6. Performing Organization Code:
7. Author(s) Alexandria I. Rossi-Alvarez Kevin Grove Charlie Klauer Melissa Miles Andy Schaudt Zac Doerzaph		8. Performing Organization Report No. VTTI-00-027
9. Performing Organization Name and Address: Safe-D National UTC Virginia Polytechnic Institute and State University		10. Work Unit No. 11. Contract or Grant No. 69A3551747115/00-027
12. Sponsoring Agency Name and Address Office of the Secretary of Transportation (OST) U.S. Department of Transportation (US DOT)		13. Type of Report and Period Final Research Report 14. Sponsoring Agency Code

15. Supplementary Notes

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.

16. Abstract

The advancement of SAE Level 4+ Automated Vehicles (L4/5 AVs) has led numerous stakeholders to develop external communication systems for these vehicles. Most research on vehicles emulating these displays has been conducted using one vehicle. However, it is vital to understand how communication to vulnerable road users (VRUs) is affected when multiple L4/5 vehicles are present. This study examined how L4/5 AVs can best communicate their intentions (e.g., turning, stopping, yielding) to VRUs and drivers of conventional vehicles. Subjective and objective data was collected to assess road user responses to two vehicles emulating L4/5 displays, from both a passenger and pedestrian perspective. Participants with no prior knowledge of the experiment's design or intent experienced three light patterns that provided information regarding L4/5 AVs' intent to slow/stop, begin, and travel with simulated automation active. Overall, participants were overwhelmed by multiple vehicles with different light bars in their crossing vicinity and found it difficult to prioritize attention. These results have implications for future design of external communication displays on L4/5 AVs. Training may be necessary for road users, given the relatively low percentage of participants who understood the meaning of these displays after multiple exposures and participants' confusion in where to look and how to interpret the intention of the displays when multiple vehicles were present.

17. Key Words Publication, guidelines, report, brochure, communication, marketing, SAE Level 4+ Automated Vehicles, external communication, SAE L4/5 AV, pedestrian, vulnerable road users, VRU, driving scenario		18. Distribution Statement No restrictions. This document is available to the public through the Safe-D National UTC website, as well as the following repositories: VTechWorks, The National Transportation Library, The Transportation Library, Volpe National Transportation Systems Center, Federal Highway Administration Research Library, and the National Technical Reports Library.	
19. Security Classif. (Of this report) Unclassified	20. Security Classif. (Of this page) Unclassified	21. No. of Pages 32	22. Price $0

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Abstract

The advancement of SAE Level 4+ Automated Vehicles (L4/5 AVs) has led numerous stakeholders to develop external communication systems for these vehicles. Most research on vehicles emulating these displays has been conducted using one vehicle. However, it is vital to understand how communication to vulnerable road users (VRUs) is affected when multiple L4/5 vehicles are present. This study examined how L4/5 AVs can best communicate their intentions (e.g., turning, stopping, yielding) to VRUs and drivers of conventional vehicles. Subjective and objective data was collected to assess road user responses to two vehicles emulating L4/5 displays, from both a passenger and pedestrian perspective. Participants with no prior knowledge of the experiment's design or intent experienced three light patterns that provided information regarding L4/5 AVs' intent to slow/stop, begin, and travel with simulated automation active. Overall, participants were overwhelmed by multiple vehicles with different light bars in their crossing vicinity and found it difficult to prioritize attention. These results have implications for future design of external communication displays on L4/5 AVs. Training may be necessary for road users, given the relatively low percentage of participants who understood the meaning of these displays after multiple exposures and participants' confusion in where to look and how to interpret the intention of the displays when multiple vehicles were present.

Acknowledgements

We would like to thank our industry partners for their support, assistance, and guidance throughout the duration of the study. Thank you to Melissa Miles from State Farm for your input on the study design, research protocols, and analysis. Thank you to John Shutko and Susana Marulanda Villa from Ford for providing one of the testing vehicles and input on the study design. Thank you to Beno Loeffler and Ralf Krause from Daimler for providing one of the vehicles used for testing and input on the study design.

This project could not have been completed without the help of many individuals at Virginia Tech Transportation Institute (VTTI). Many thanks to Yubin Hong, Josh Radlbeck, Kathryn Meissner, Andres Coello, and Mario Jones, whose help was instrumental to run the experiment. Thank you to Richard Parks and Tia Farese for patiently coding the crossing-decision data. Thank you to all the different teams at VTTI for their time, efforts, and support.

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.

Table of Contents

List of Figures

List of Tables

Introduction

Drivers communicate with other road users in a myriad of ways, including turn signals, trajectory of vehicle, facial expressions, and gestures. With the introduction of different levels of automated vehicles (AVs) onto our roadways, it is essential to understand the interaction and relationships among surrounding vehicles and vulnerable road users (VRUs). Specifically, we need to understand communication between drivers, VRUs, and AVs (SAE International, 2021). This will allow engineers and designers to create systems that facilitate safe and intuitive interaction. Accomplishing this will require an understanding of the principles and fundamentals that make up the interactions of communication and AV system design.

While SAE Level 4+ AVs (L4/5 AVs) have yet to be fully integrated into the vehicle fleet, communication between drivers and VRUs is not a new challenge. Some human operated vehicles have been communicating intent to other drivers and VRUs since 1925 (Habibovic et al., 2018). This includes formal communication (e.g., turn signals, brake lights) and informal communication (e.g., eye contact, hand waving). When a driver needs to make a right turn, they need to inform outside users of their intention. To do this, they turn on the right blinker, and the vehicle uses external communication via a turn signal in the front and back of the vehicle to indicate intent and change in directional path. The vehicle also uses internal communication to confirm to the driver that the vehicle completed its task through a visual indicator on the dashboard and an audible signal. These forms of communication establish relationships between entities that share a flow of information.

As L4/5 AVs are integrated into a mixed traffic environment, communication channels need to be created to produce a connection between these AVs and surrounding road users (Schieben et al., 2019). However, there is currently a lack of knowledge about how road users will interpret an L4/5 AV's intent and/or interact with multiple L4/5 AVs in a natural setting. Automated driving systems are under development and being tested on public roadways throughout the country. Thus, it is imperative that human factors and traffic safety researchers understand and identify effective communication strategies between AVs and road users that enhance safety within the transportation ecosystem as AVs are deployed.

Additional forms of formal external communication are being designed and tested to better understand L4/5 AV and road user interactions. Specifically, work is being done to determine if these different methods aid in perception and communication of intent and if they are reliable throughout varied interactions. A study found that passengers riding in L4/5 AVs felt that additional information regarding the vehicle's intent was essential for developing passenger trust and comfort. The study looked specifically at turn signals, navigation, and hazard detection (Basantis, 2019).

The investigation of potential L4/5 AV external communication systems has been undertaken by numerous stakeholders in many countries. While some L4/5 AV external communication systems show promise, there remain design parameters that stakeholders cannot agree on, and the implications of these disagreements could negatively impact overall safety or the potential for a standards consensus. The two most critical areas of contention are currently (1) the best location to position a visual external communication signaling system and (2) the color of the visual signaling system. These two parameters affect the salience of the communication devices as well as interpretation. For example, road users may have mental models for vehicle lights (i.e., emergency vehicles, service vehicles), which may also impact understanding.

For this project, LED strips mounted on the front of a vehicle (see Figure 1) were used based on the literature review. Faas et al. (2020) found that participants felt safer, had more favorable experience ratings, and perceived a highly automated vehicle (HAV; equivalent to an L4/5 AV) as intelligent and transparent when any external human–machine interface (eHMI) was present. In addition, informing pedestrians about the HAV’s intent improved their sense of safety (Faas et al., 2020). Audio was not included in the eHMI because researchers believed it might be difficult for pedestrians to pick up on these sounds in busy traffic environments (Lee et al., 2019), a condition that we aimed to emulate.

Most research on the usage of external communication with L4/5 AVs has involved only one L4/5 AV within the environment. The research that has included more than one L4/5 AV was conducted in virtual or augmented environments. This has repercussions in terms of external validity and understanding how road users will interact in real-world settings. Including two or more L4/5 AVs in testing scenarios will allow for insights on how VRUs’ perspectives and behaviors are affected when introduced to more complex traffic scenarios. Additionally, it will provide a better understanding of how interacting with multiple L4/5 AVs impacts VRUs’ willingness to cross the street in close proximity to these vehicles or any confusion that results regarding the intent of multiple L4/5 AVs.

This study examined how vehicles emulating an SAE L4/5 can best communicate with other drivers and VRUs about their intentions (e.g., turning, stopping, yielding, etc.). These different external communication methods were evaluated through varied location and color design parameters. Dynamic scenarios on a test track were used to assess interaction and decision-making. Additionally, multiple L4/5 AVs were included in each testing condition to gain a deeper understanding about complex roadway scenario interactions (i.e., a four-way stop, construction zone) and how that impacts VRUs’ interpretation of vehicle intent.

Research Questions

This research examined how an L4/5 AV can best communicate with other drivers and VRUs about its intentions (e.g., turning, stopping, yielding, etc.). Specifically, this study sought to understand where the most effective location for an external visual communication system should be located on an L4/5 AV as well as the most effective color to use for signaling when communicating with other drivers and pedestrians. Additionally, we wanted to know if the participants were able to learn the meaning of the patterns over time. Several key research questions were addressed.

External Communication Impact

- Does glance and decision-making behavior change once a participant understands the external communication displays?
- How many exposures to vehicle’s external communication does it take for participants to understand the meaning of the displays?

Light Bar Factors

- Does the location of the vehicle's external communication influence glance behavior and decision-making?
- Does the color of vehicle's external communication influence glance and decision-making behavior?
- Is there an influence on glance or decision-making behavior when the L4 AV is still in motion or stopped?
- How many exposures to the light bar patterns does it take until participants correctly interpret the meaning of all three intentions?

Testing Scenarios

- How are VRUs' perspectives and behaviors affected when they are introduced to more complex traffic scenarios?
- Does the objective data collected across these scenarios correlate to the qualitative information collected?
- Do these testing scenarios provide reliable human performance data, specifically in measures of decision-making and glance?

Method

Study Design

To study participants' understanding of vehicle intention with the L4/5 AVs, a within-subject design was used. Forty participants observed external communication displays on two vehicles that emulated L4/5 AVs. Every participant viewed eight different testing scenarios that were repeated for multiple trials. The eight scenarios were split between pedestrian scenarios (where the participants assumed the role of pedestrian on the roadside) and passenger scenarios (where the participants assumed the role of a passenger in a non-automated vehicle). When participants were experiencing the pedestrian scenarios, they would make crossing decisions and complete rankings on their understanding of the displays. When participants were experiencing the passenger scenarios, they would observe the scenario and complete ratings on their understanding.

Procedure

After participants arrived at the testing facility and were cleared through the COVID-19 protocols, they were led to a conference room for the pre-session paperwork. They were asked to complete the Informed Consent Form, a W-9 (for compensation purposes), and a pre-session questionnaire. Next, participants were given a baseline visual acuity test using a Snellen eye-exam chart. They were also assessed for color blindness using the Ishihara Colorblind Assessment. Participants had to have a minimum visual acuity of 20/40 (corrected to normal vision) to continue. A participant could continue if they were colorblind; however, the information was noted to account for any outliers.

Upon arrival at the intersection, the moderator provided instructions on the safety procedures and decision-making box. To reinforce the façade that the test vehicles were driverless L4/5 AVs, the moderator communicated with the control tower to program the vehicles for a specific scenario and trial. After the participant was exposed to the trial, they completed a series of questions about their experience. The moderator provided verbal probes to understand the participants' experience. This process was repeated until participants had experienced all scenarios and all trials.

After participants experienced all scenarios on the Virginia Smart Roads Surface Streets, they proceeded back to the conference room. Then the moderator debriefed participants regarding the need for deceit about the "fully automated vehicle." Moderators explained that it was important for participants to believe that the vehicles were highly automated to ensure that their perceptions and responses regarding their decisions to cross the streets would generalize to traffic scenarios where an automated system may control vehicles.

Variables

There were several conditions included in the study. The independent variables are depicted in Table 1.

Table 1. Independent Variables

VARIABLE	LEVELS	DESCRIPTION
LIGHT BAR LOCATION	2	1) Grill 2) Windshield
LIGHT BAR COLOR	3	1) White Thick 2) White Thin 3) Teal Thin
LIGHT BAR PATTERN	3	1) Yield 2) Stop 3) Proceed
PEDESTRIAN TESTING SCENARIOS	4	1) Scenario 1 (4x) 2) Scenario 2 (2x) 3) Scenario 3 (2x) 4) Scenario 4 (2x)
PASSENGER TESTING SCENARIOS	4	5) Scenario 5 (2x) 6) Scenario 6 (2x) 7) Scenario 7 (2x) 8) Scenario 8 (2x)

Dependent variables included several forms of measurement.

- **Crossing Decision**: Quantify how many times the participant decided to cross or not cross the street across scenario and vehicle condition.
- **Motion of L4/5 AV**: The number of crossing decisions prior to the vehicle coming to a full stop.
- **Glances to L4/5 AV**: The number of glances the participants made to the L4/5 AV throughout a single scenario trial.
- **Learning Over Exposure**: Record the number of trials required before the participant correctly interpreted the patterns of the light bars.

Participant Demographics

Forty participants were recruited from the New River Valley using the Virginia Tech Transportation Institute (VTTI) participant database. Volunteers were between 18 to 55 years old (M = 38, SD =11.82), and evenly distributed by gender (male = 20, female = 20). Participants were compensated $60.00 for completing the two-hour session. This research protocol was approved by the Virginia Tech Institutional Review Board (Protocol # 20-790).

Study Location

The study was conducted in VTTI meeting rooms and on the Virginia Smart Roads Surface Street. The Virginia Smart Roads are closed test-bed research facilities managed by VTTI. Only vehicles and personnel associated with the study were allowed access to the track during testing.

Testing L4/5 AVs

Two pseudo L4/5 AVs were used in this study. Both vehicles were retrofitted with mounted LED light bars in the grill and the top of the windshield. Both LED bars were controlled by a switch mounted inside the vehicle. Testing vehicles were always at least one car lane away from the participant during pedestrian scenarios. AV-B was equipped with automatic emergency braking, which was active during testing. Both vehicles were equipped with a data acquisition system (DAS), which included cameras recording the forward view to capture participant behavior as a pedestrian. Vehicles did not have their headlight active during the study. All vehicles followed standard traffic laws and utilized their turn signals when necessary.

L4/5 AVs' External Interface Design

There were two emulated L4/5 AVs used in this study: AV-A only displayed white LED light bars, whereas AV-B's LED light bars were changed between the colors white and teal (shown in Figure 1; the vehicles will be referred to AV-A and AV-B for the entirety of the report). The light bars' location (i.e., windshield or grill) and color were changed between scenarios, so all participants experienced both locations and all color combinations. This was counterbalanced across participants to minimize any order effects. For a baseline measurement, the light bars were turned off for two trials in scenario 1.

Figure 1. Overview of L4/5 AV's external communication location and color.

Both L4/5 AVs communicated vehicle intent using three different states with the pedestrian: (1) Drive state, (2) Yield state, (3) Ready to drive state. The Drive state was communicated by the light bar being uniformly lit (i.e., no motion or blinking) and indicated that the vehicle was in motion and would remain in motion. The Yield state was communicated by the light bar flashing outwards to inwards and indicated that the vehicle was preparing to stop. The Ready state was communicated by the light bar blinking three times and indicated that the vehicle was about to start moving again from a stopped position.

All three of the vehicle intent light patterns (i.e., Drive, Yield, Ready) were displayed by each vehicle in all scenarios and trials. However, the display time of the Drive and Yield states was naturally more prolonged than the Ready state, providing participants more time to view the pattern. In addition, the Drive and Yield pattern exposure were naturally longer because of the distance the vehicle was driving, and the length of time the "AV" potentially stopped was longer than when the AV indicated the Ready pattern. This exposure difference matches what would happen in real-world scenarios.

Participants had no prior knowledge of the lights or their location, pattern, or color. Participants were never taught the meaning of the different patterns throughout the entirety of the session. This was to understand participants' learning of the existence and meaning of the lightbars over number of exposures to the systems.

The study sponsors procured all testing vehicles and arrived equipped with their own light bars. AV-A's light bar was slightly thicker than the light bar installed in AV-B. However, prior to testing, the luminance levels were equalized for both vehicles' light bars.

Light Bars Luminance Equalization

For the two different models of vehicles utilized, there were different lighting implementations that resulted in two different light levels and distributions. AV-B utilized diffusion optics on white LEDs while AV-A utilized semi-cylindrical optics on LEDs, which resulted in different distribution. To measure the differences, a luminance meter on a tripod mounted with the lens was utilized; this was 1 meter tall and 2 meters radially away from the lighting elements on each vehicle. Measurements were then made on azimuthal intervals from 0 ° (directly in front of the vehicle) to 90 ° or nearly 90 ° (perpendicular to the direction of travel). Results showed that AV-B had a diffuse distribution versus AV-A's more forward distribution.

To address the differences and reduce the number of factors between vehicles, diffusion optics were 3D printed and mounted in front of the AV-A optics. The lighting level of AV-A was then adjusted by changing the voltage so that the grill lighting was at the same overall lighting level within 1% average over the 0–67.6 ° azimuthal range (Figure 2, left). The lighting above the windshield required attenuation of AV-B's output, so a 0.3 neutral density filter was placed over the optics. The filter resulted in the output shown in Figure 2 (right) and resulted in lighting within 15% average over the range of 0–67.5 °.

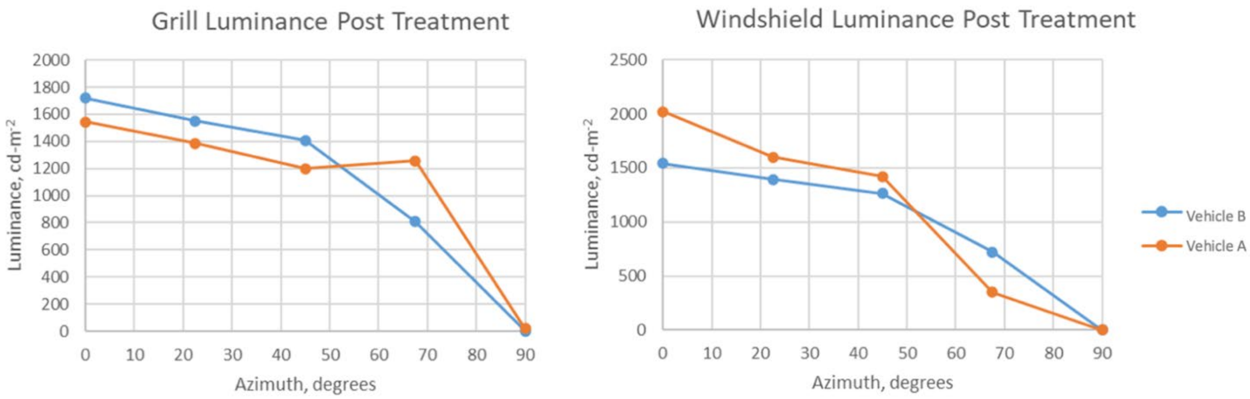

Figure 2. Lighting level and distribution differences between AV-A and AV-B at the grill (left) and above the windshield (right) post treatment.

Emulation of L4/5 AVs

At the beginning of the study, participants were informed that there would be two L4/5 AVs on the test track. The moderator did not expressly state that there were no drivers present. The operators were wearing a seat-suit (Figure 3), which allows a human driver to be disguised as an empty driver's seat, thus creating the illusion of a fully autonomous vehicle. This was necessary to test and evaluate real-world encounters and behaviors. The drivers wearing the seat-suit were verified to have an adequate field of vision to ensure the suit did not create any potential safety hazards.

Seat-Suit A Seat-Suit B

Figure 3. Seat-suit used for each L4/5 AV.

Testing Scenarios

All scenarios chosen were common traffic scenarios that may cause confusion for humans interacting with L4/5 AVs. All participants experienced eight different scenarios (four as pedestrians, four as passengers) as depicted in Table 2 within a 2-hour session. The first scenario was repeated four times, and for two of these four trials one vehicle had no external communication display (baseline condition). All other scenarios were repeated two times (all with lighted external communication displays). Participants experienced a total of 18 testing trials. The light bar color and location signal changed across each repetition of the scenario, which was counterbalanced to mitigate any order effects.

For the pedestrian scenarios, the crossing path and number of vehicles impacting that path was determined and used throughout the analysis. Scenario 1 had one vehicle with a straight crossing path (1 Veh SCP) in front of pedestrian. Scenario 2 included a right turn (RT). Scenario 3 had two vehicles that impeded on pedestrians' straight crossing path (2 Veh SCP). Details and illustrations of each scenario are included in Appendix A.

Table 2. Testing Scenario Matrix

Scenario Type	Scenario	L4/5 AV Crossing (Only Pedestrian)	Light Bar Presence	Data Collected	Speed	Number of Trials
Pedestrian	4-way stop	1 Vehicle	Vehicle A – Light bars on for 4 trials Vehicle B – Light bars on for *only* 2 trials	Crossing-decision, Glance, Distance, Qualitative feedback, Survey Data	10 mph	4
Pedestrian	Vehicle Right Turn at a 4-way stop	2 Vehicles	Vehicle A and B – Light bars on for all trials	Crossing-decision, Glance, Distance, Qualitative feedback, Survey Data	10 mph	2
Pedestrian	Both Vehicles Proceed Straight	2 Vehicles	Vehicle A and B – Light bars on for all trials	Crossing-decision, Glance, Distance, Qualitative feedback, Survey Data	10 mph	2
Pedestrian	Lane Change and Simulated Pedestrian	2 Vehicles	Vehicle A and B – Light bars on for all trials	Go/no-go decision, Qualitative feedback, Survey Data	35 mph	2
Passenger	Lane Change with Expert Pedestrian	N/A	Vehicle A and B – Light bars on for all trials	Qualitative feedback, Survey Data	10 mph	2
Passenger	4-way Stop and Vehicle Left Turn	N/A	Vehicle A and B – Light bars on for all trials	Qualitative feedback, Survey Data	10 mph	2
Passenger	Construction Zone with Active Flagger	N/A	Vehicle A and B – Light bars on for all trials	Qualitative feedback, Survey Data	10 mph	2
Passenger	Mid-block Lane Change	N/A	Vehicle A and B – Light bars on for all trials	Qualitative feedback, Survey Data	35 mph	2

Due to the complexity of the scenarios, the scenario order did not change. However, to mitigate any order effects, half of the participants (n = 20) experienced the pedestrian scenarios first followed by the passenger scenarios, and the other half of participants (n = 20) experienced the passenger scenarios first followed by pedestrian scenarios.

Decision- Box – To Cross the Street

To comply with safety protocols, participants were not allowed to cross the street in front of the testing vehicle; a decision box was utilized instead (see Figure 4). Participants stood outside the box as the moderator provided instructions. Participants were asked to decide when they deemed it was safe to cross the street without physically crossing the street. When they felt that they would cross the street, they stepped inside the box. When they felt that they would not cross the street, they stepped back outside the box. They were allowed to step back and forth inside or outside the box as many times they wanted per trial.

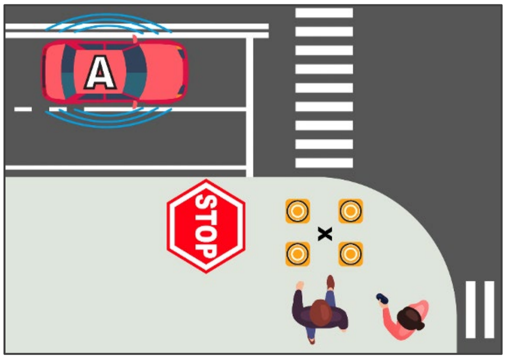

Figure 4. Aerial depiction of the decision-making box.

Results and Discussion

Analysis Overview

Participant video data was collected from each DAS installed in the L4/5 AVs. The DAS captured all kinematic and driver performance data and video of the participants' decision-making. The video and kinematic data was combined to understand vehicle distance and speed compared to a participant's decision to cross or not cross the street. Due to limitations with the camera clarity, movements of the head, as opposed to the eyes, were used to define glance direction for coding.

A primary coder for each vehicle measured the decision to cross, glance, distance of decision, vehicle movement, duration of decision, and indecision. A second coder analyzed a subset of data from each vehicle. A Pearson product-moment correlation was run to determine interrater reliability. There was a strong interrater reliability between the two coders ($r = .486$, $n = 70$, $p = .001$). As the correlation coefficient is closer to +1, this indicated a positive association, meaning the coders' data varied (either increased or decreased) in the same direction.

Crossing Decision Analysis

Crossing decisions were measured by the number of times a participant decided to cross the street during an active scenario (both vehicles were in motion completing the scenario); i.e., the number of times they fully stepped inside the decision- box. This data was analyzed across the scenario trial, condition tested, and vehicle movement.

Willingness and Unwillingness to Cross by Scenario

Forty pedestrians were asked to utilize the decision-making box to make a crossing decision for three different traffic scenarios where L4/5 AVs were operational. Cochran's Q test was run to determine if the percentage willingness or unwillingness to cross was different across the three different testing scenarios. Across all trials for 1 Veh SCP, 82.3% of pedestrians were willing to cross. Across all trials for RT, 72.6% of pedestrians were willing to cross. Across all trials for 2 Veh SCP, 83.9% of pedestrians were willing to cross. The percentage of pedestrians willing to cross the street was statistically significantly different across the different scenarios, $\chi^2(2) = 7.167$, $p < .05$.

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. There was a statistically significant increase in the percentage of crossing from RT to 2 Veh SCP ($p < .05$). There was no statistically significant difference in the willingness to cross percentage from RT to 1 Veh SCP, and 1 Veh SCP to 2 Veh SCP.

Willingness and Unwillingness to Cross by Trial

Willingness and unwillingness to cross was analyzed across eight trials that utilized the decision-making box. Figure 5 outlines the frequencies of participants' willingness or unwillingness to cross the street for each trial. The highest crossing frequency was for 1 Veh SCP trial 3 (1 Veh SCP) and 2 Veh SCP trial 2 (2 Veh SCP).

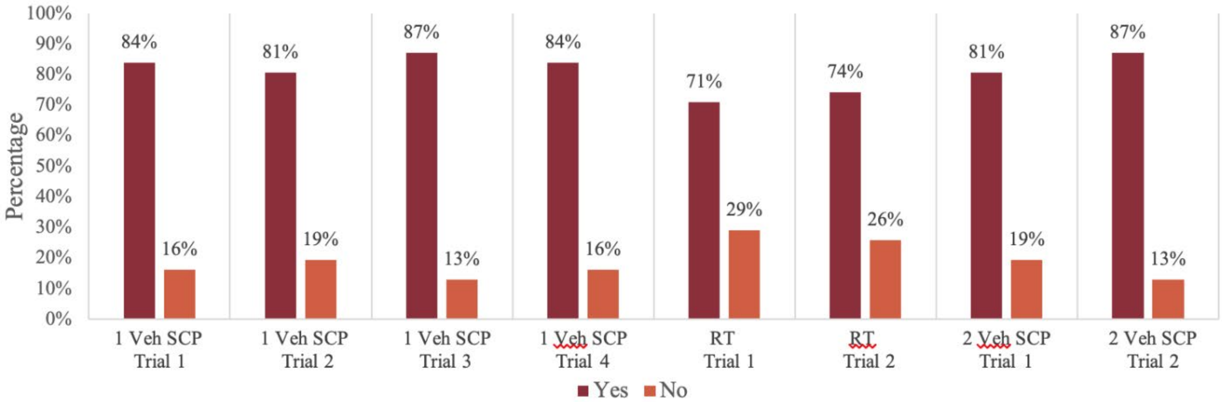

Figure 5. Frequencies of willingness and unwillingness to cross across scenario and trial.

Cochran's Q test was used to assess the difference between the testing scenarios and participants' decision to cross. The sample size met assumptions, so the χ^2-distribution approximation was used. The percentage of participants who were willing to cross the street was not statistically significantly different across trials, $\chi^2(7) = 13.21$, $p = .067$.

Number of Crossing Decisions by Scenario

The number of times a participant decided to cross was coded. A one-way repeated measures analysis of variance (ANOVA) was conducted to determine whether there were statistically significant differences in the number of crossings pedestrians made across each scenario type. There were no outliers, and the data was normally distributed, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively. Mauchly's test of sphericity indicated that the assumption of sphericity had been violated, $\chi^2(2) = 13.257$, $p < .001$. Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.835$). The number of times pedestrians decided to cross the street was statistically significantly different across the scenario types, $F(2, 122) = 5.566$, $p < .05$, partial $\eta^2 = .084$. Post hoc analysis with a Bonferroni adjustment revealed that 1 Veh SCP and 2 Veh SCP both have a higher crossing decision percentage than RT scenarios.

Number of Crossing Decisions by Trial

The number of times a participant decided to cross was coded. A one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences in the number of crossings pedestrians made across eight trials.

There were no outliers, and the data was normally distributed, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(27) = 57.252$, $p < .001$. Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.573$). The individual trials elicited statistically significant changes over number of crossings, $F(7, 210) = 3.054$, $p < .05$, partial $\eta^2 = .092$.

Post hoc analysis with a Bonferroni adjustment revealed that the number of crossings only significantly decreased from trial 4 to trial 5 (M = 0.419, 95% CI [0.072, 0.767], $p < .05$), and no other trials.

Crossing Decision During Vehicle Approach

Participants' decisions to cross the street when the L4/5 AV was approaching the intersection (i.e., slowing to a stop near stop sign controlled intersection) is depicted in Figure 6. Cochran's Q test was used to assess the difference between the testing scenarios and decision to cross when the L4/5 AV was moving. The percentage of participants who crossed when the L4/5 AV was still moving was not statistically significantly different across trials, $\chi^2(7) = 2.33$, $p = .939$.

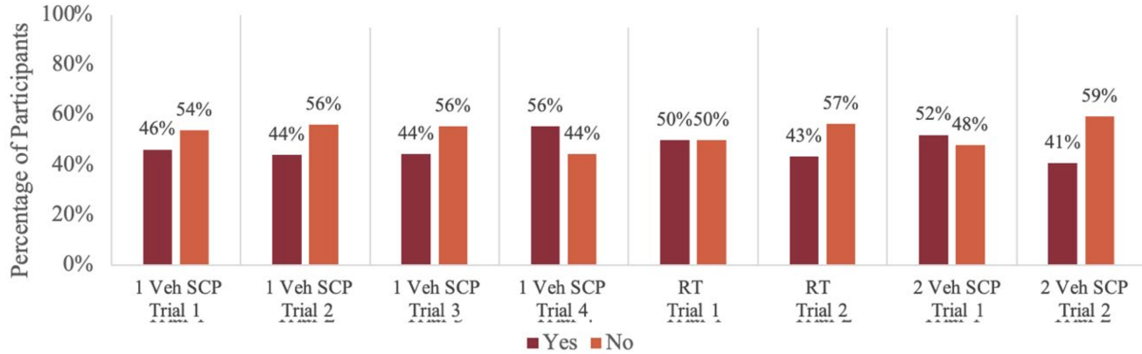

Figure 6. Frequencies of crossing decision when vehicle in motion.

Crossing Decision Across Light Bar Condition

Differences Between Condition by Willingness or Unwillingness to Cross

Crossing decision by light bar condition percentages are indicated in Figure 7. Among the conditions, the thin white windshield had the highest crossing percentage.

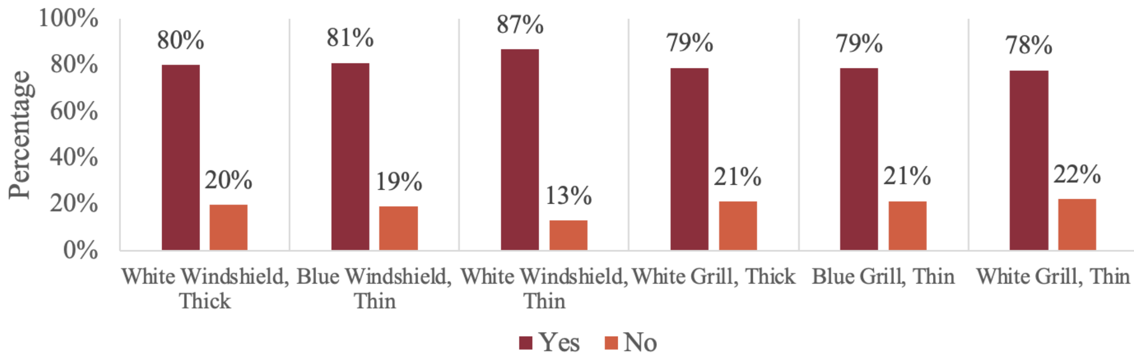

Figure 7. Frequencies of crossing decision by light bar condition.

Cochran's Q test was used to assess the difference between the light bar conditions and participants' decisions to cross. The sample size met assumptions, so the χ^2-distribution approximation was used. The percentage of participants who were willing to cross the street was not statistically significantly different across different light bar conditions, $\chi^2(5) = 7.11$, $p = .213$.

Glances to L4/5 AV Analysis

Glances to Both L4/5 AVs by Scenario

The total number of glances participants made to the L4/5 AV was coded across each scenario to understand which vehicle participants were attending to. A one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences across the three scenarios and the number of pedestrian glances to the L4/5 AVs. There were no outliers, and the data was normally distributed, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 12.141$, $p < .001$. Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.845$). The difference scenarios elicited statistically significant changes in the number of glances, $F(2, 122) = 73.680$, $p < .001$, partial $\eta^2 = 0.547$. Post hoc analysis with a Bonferroni adjustment revealed that the number of glances statistically significantly decreased from 1 Veh SCP to RT (M = 1.145, 95% CI [0.831, 1.460], $p < .001$), from 2 Veh SCP to 1 Veh SCP (M = .565, 95% CI [0.142, 0.987], $p = .005$), and from 2 Veh SCP to RT (M = 1.710, 95% CI [1.398, 2.022], $p < .001$).

Glances to Both L4/5 AVs by Trial

The total number of glances participants made to the L4/5 AV was coded across each scenario and trial, as shown in Figure 8. A one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences across the eight trials and the number of glances to the L4/5 AVs. There were no outliers, and the data was normally distributed, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(27) = 65.714$, $p < .001$. Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.701$).

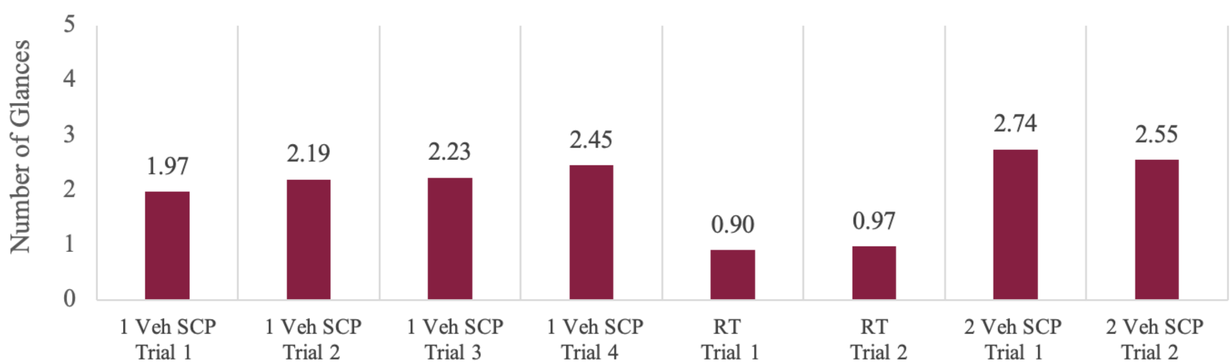

Figure 8. Number of glances to both vehicles across scenario and trial.

The difference trials elicited statistically significant changes in the number of glances, $F(7, 210) = 19.391$, $p < .001$, partial $\eta^2 = 0.393$. Post hoc analysis with a Bonferroni adjustment revealed that the number of glances was statistically significantly different across several trials.

Post hoc analysis with a Bonferroni adjustment revealed there were statistically significant pairwise comparisons; this is illustrated in Appendix B. RT Trial 1 was significantly different from all other trials, except trial RT Trial 2. RT Trial 2 was significantly different from all other trials, except RT Trial 1. 2 Vehicle SCP Trial 1 was statistically significant, decreasing from 1 Vehicle SCP Trial 1 (M = .774, 95% CI [7.651, 1.548], $p = .050$).

Number of Glances by Individual L4/5 AVs

The total number of glances participants made to each specific L4/5 AV (A or B) throughout each scenario trial was coded, as shown in Figure 9.

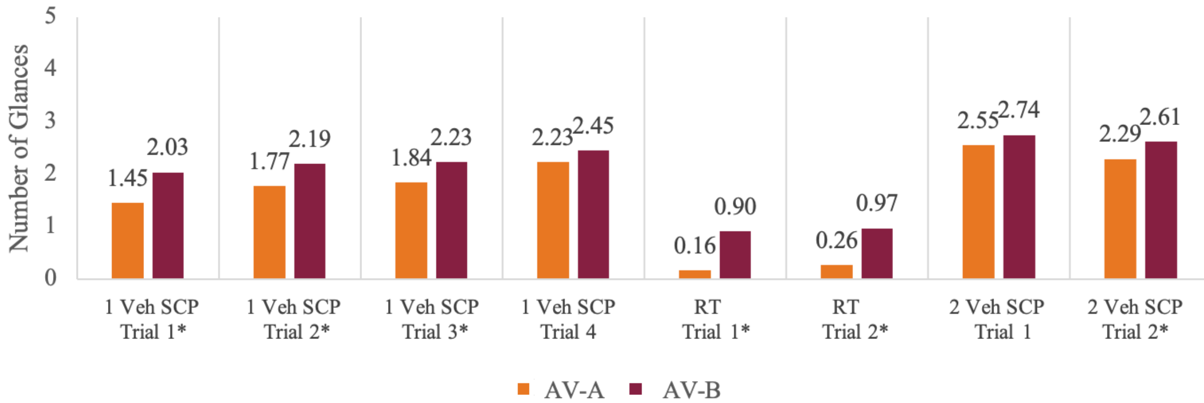

Figure 9. Number of glances for each vehicle across scenario and trial.

A two-way repeated measures ANOVA was run to determine how the number of glances was affected by the two vehicles across the difference trials. Analysis of the studentized residuals showed that there was normality, as assessed by the Shapiro-Wilk test of normality and that there were no outliers, as assessed by no studentized residuals greater than ± 3 standard deviations. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the interaction term, $\chi^2(27) = 29.774$, $p = .330$. There was a statistically significant interaction between vehicle and trial on the number of glances, $F(7, 220) = 3.486$, $p = .001$, partial $\eta^2 = .104$. Therefore, simple main effects were run (see Table 3).

Table 3. Glance By Individual L4/5 AV's Pairwise Comparison Results

Trial Comparison	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference Lower Bound	95% Confidence Interval for Difference Upper Bound
1 Veh SCP, Trial 1	.581*	.121	<.001	.334	.827
1 Veh SCP, Trial 2	.419*	.137	.005	.139	.700
1 Veh SCP, Trial 3	.387*	.110	.001	.161	.613
1 Veh SCP, Trial 4	.226	.111	.050	.000	.452
RT, Trial 1	.742*	.092	<.001	.553	.931
RT, Trial 2	.710*	.083	<.001	.540	.879
2 Veh SCP, Trial 1	.194	.097	.056	-.006	.393
2 Veh SCP, Trial 2	.323*	.142	.031	.032	.613

Total Number of Glances by Crossing Path

The scenario type was condensed to scenarios that had one vehicle crossing the pedestrians' path, and where both vehicles were crossing their path, as shown in Figure 10.

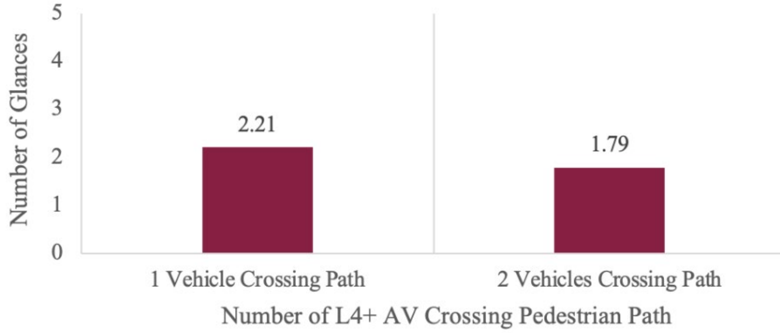

Figure 10. Number of glances to both vehicles across scenario and trial.

A paired-samples t-test was used to determine whether there was a statistically significant mean difference between the number of glances across number of vehicles crossing a participants' path. There were no outliers in the data, as assessed by inspection of a boxplot. Participants had a higher average of glances for 1 Veh SCP (M = 2.209, SD = 1.068) as opposed to 2 Veh SCP (M = 1.790, SD = 1.106). This could have been because they were fixated on one vehicle that they self-determined to be the larger threat to their safety when crossing the street. The 1 Veh SCP elicited a statistically significant increase of .419, 95% CI [0.159, 0.678] in number of glances compared to 2 Veh SCP, $t(123) = 3.198$, $p = .002$, d = .287.

Learning Over Exposure Analysis

After each exposure to the light bars, participants' knowledge of the light bar patterns and correct interpretation of the patterns was coded. The following sections analyze how long it took participants to notice the light bar pattern, and correctly articulate the pattern meaning.

Analysis of Noticing the Patterns

Forty participants were exposed to the light bar patterns over 18 trials, and their indication that they noticed the light bars was denoted by the experimenter during the semi-structured interviews after each exposure (see Figure 11).

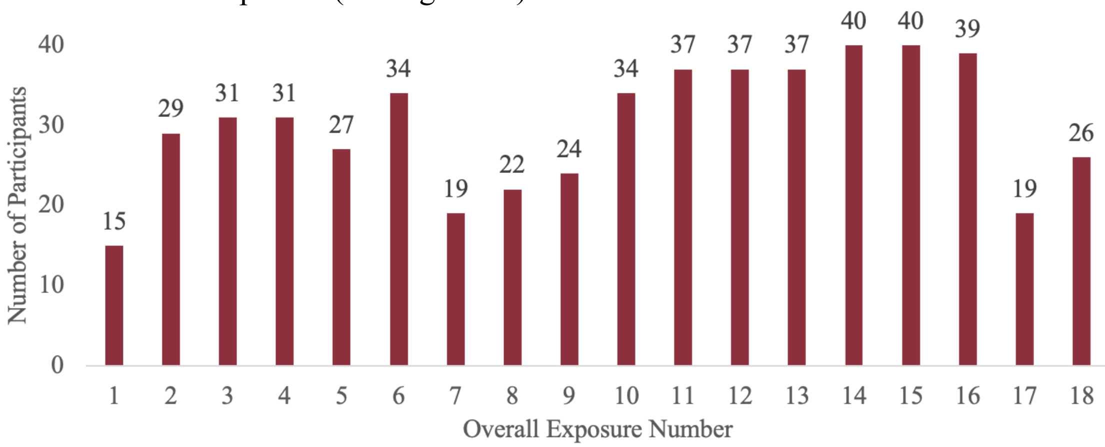

Figure 11. Participants who noticed the light bars over exposure.

Cochran's Q test was run to determine if the percentage of participants noticing the light bars was different at the different time points. Sample size was satisfactory to use the χ^2-distribution approximation. The percentage of participants noticing the light bars was statistically significantly different at the different time points, $\chi^2(17) = 174.629$, $p < .001$. Pairwise comparisons were completed using Dunn's procedure with a Bonferroni correction for multiple comparisons. Exposure 1, 9, 17, and 18 had statistically significantly pairwise comparisons; this is illustrated in Appendix C.

Analysis of Correctly Interpreting Patterns

The number of trials participants took to identify the meaning of all patterns completely and correctly was documented (see Figure 12). Cochran's Q test was run to determine if the percentage of participants understanding the patterns was different across the number of exposures. Sample size was satisfactory to use the χ^2-distribution approximation. The percentage of learning was statistically significantly different over number of exposures, $\chi^2(17) = 79.928$, $p < .001$. Pairwise comparisons were completed using Dunn's procedure with a Bonferroni correction for multiple comparisons resulting in 16 significantly different comparisons (see Appendix D).

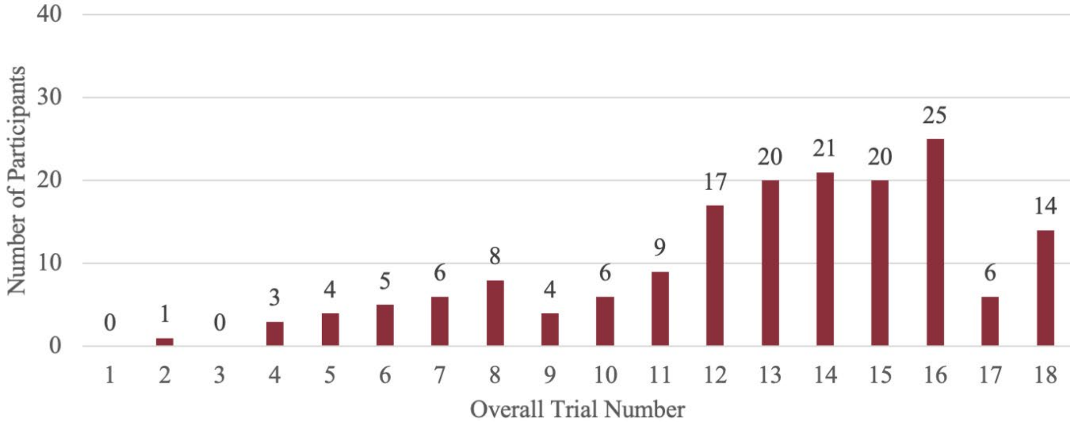

Figure 12. Number of exposures until correct interpretation across all patterns.

Qualitative Analysis of Pattern Understanding

All participants experienced all three light patterns across all scenarios and trials. Participants were never provided any information on the patterns. Participants were noted as successfully understanding the patterns when they could discern all three patterns and their specific purpose.

It took participants more than 12 exposures to the light bars to begin to understand what the patterns meant. After 16 exposures, the largest percentage of participants understood all patterns. However, at least 15 participants never understood what the patterns represented.

Yield and Driver patterns were most recognized over the Ready pattern. Yield and Driver patterns were the first two patterns that participants were able to interpret. Some participants only noticed the first two patterns across the entire study. The Ready pattern was the most difficult and last pattern participants understood. This is likely because the pattern length was too short for participants to make note of.

Participants stated that it was difficult to watch and interpret the light patterns for two vehicles in their environment. Often, they had to prioritize their focus on the vehicle they felt had the most risk to their crossing decision. They also missed the light patterns on the vehicles because they turned their heads back and forth to look at the other vehicles. During this movement, they missed the short Ready pattern change.

Light Bar Desirability

After participants completed all testing scenarios and trials, they were asked to rank all six light bar conditions they experienced on a scale from 0 (least desired) to 50 (most desired). Results are shown in Figure 13. Through this method, researchers were able to gain both qualitative feedback about their preferences and quantitative results.

Desirability Quantitative Findings

Figure 13. Desirability scale rating per condition.

A Friedman test was run to determine if there were differences among the different light bar conditions across the ranked desirability scores. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. Light bar conditions were statistically significantly different across varied conditions on the desirability scale, $\chi2(5) = 45.733$, $p < .001$. Post hoc analysis revealed statistically significant differences, as indicated in Table 4 (sample 1 was more preferred than sample 2).

Table 4. Pairwise Comparison of Light Bar Condition Across Desirability Ratings

Sample 1 - Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.
Thin White Grill - Blue Grill	0.463	0.418	1.106	0.269
Thick White Grill - Thin White Grill *	-1.225	0.418	-2.928	0.003
Thin White Windshield - Thin White Grill *	-1.425	0.418	-3.406	< .001
Blue Windshield - Thin White Grill *	1.775	0.418	4.243	< .001
Thick White Windshield - Thin White Grill *	-2.463	0.418	-5.887	< .001
Thick White Grill - Blue Grill	-0.763	0.418	-1.823	0.068
Thin White Windshield - Blue Grill	-0.963	0.418	-2.301	0.021
Blue Windshield - Blue Grill *	-1.313	0.418	-3.137	0.002
Thick White Windshield - Blue Grill *	-2.000	0.418	-4.781	< .001
Thin White Windshield - Thick White Grill	-0.200	0.418	-0.478	0.633
Blue Windshield - Thick White Grill	0.550	0.418	1.315	0.189
Thick White Windshield - Thick White Grill*	-1.238	0.418	-2.958	0.003
Blue Windshield - Thin White Windshield	0.350	0.418	0.837	0.403
Thick White Windshield - Thin White Windshield	1.038	0.418	2.480	0.013
Thick White Windshield - Blue Windshield	-0.688	0.418	-1.643	0.100

Desirability Qualitative Findings

Overall, the windshield conditions were clustered higher on the desirability scale versus the grill conditions. The higher placement allowed for easier visibility, and this location is where participants look for vehicle information as a pedestrian. Additionally, for both location segments, the color order was the same. Participants preferred the thicker white light bar, followed by the teal light bar, then the thinner white light bar. This preference was due to the ease of visibility of the light bar across exposures. As stated in the method section, the white display had a thick and thin component while the teal display did not.

White Windshield, Thick

Participants preferred the white windshield location the most because the light bar was thicker, wide, bold, and more prominent on the vehicle. Additionally, the location in the windshield was ideal and easier to see because it was in participants' line of sight. This condition was some participants' favorite because it was able to communicate information effectively. Participants liked that the light bar stood out because it got their attention. These participants stated that they preferred the individual lights that were segmented rather than one solid light.

Teal Windshield

Overall, participants preferred this condition due to the location of the light bar in the windshield. The location was easy to see because of placement. Some participants liked the teal color of the light bar because it contrasted the vehicle, and the teal could not be confused with the headlights. However, some participants felt this light bar condition had poor visibility. Specifically, the color and thinness of the light bar made it difficult to see.

White Windshield, Thin

The white windshield condition was the least preferred among the windshield options. Specifically, some participants thought the combination of the thinner bar with the white color made visibility more challenging. Participants felt this light bar condition blended into the vehicle (specifically the vehicle's roofline), which attributed to poor visibility. Only a handful explicitly stated the sun made it more difficult to view. However, since the light bar was placed in the windshield it was still rated higher than other conditions. Participants could see the light bar better because it was more visible on top compared to the grill location, and there were no other distracting lights.

White Grill, Thick

Some participants liked this condition because it was more visible than thin white and teal grill light bars. This light bar was more prominent due to the thickness, making it more noticeable and visible among other lights in the vicinity. However, the location of the light bar in the grill made it challenging to see because participants do not typically look for information in the grill, making it an unintuitive location for participants to look for information that would influence their crossing behavior.

Teal Grill

Most participants had a negative experience with this condition because the teal grill light bar was difficult to see. A few stated they would ignore this light bar because of the poor visibility. The teal color and grill location of the light bar negatively affected its visibility. Some participants stated that the surrounding chrome in the grill made it difficult to view the grill. A few could not determine if the light bar was an accent light or a signal. However, this was not the least desired experience because the teal color stood out a little more compared to the thin white grill light bar.

White Grill, Thin

Overall, most participants did not have a positive experience with this condition. Almost all participants did not see the light bar at one point. Some had difficulty deciphering if the light bar was on. Other participants stated the sun or the angle at which they viewed the light bar impacted the visibility. The location of the light bar in the grill negatively impacted the participants' visual perception. The low placement of the light bar in the grill and the fact that participants do not generally look at the grill for crossing information contributed to negative ratings. The light bar also blended in with the other grill components or blended in with the front headlights. The combination of the white color and the grill location was ineffective for communicating information. A few participants thought the light bar itself was too thin, making visibility difficult.

Conclusions and Recommendations

This study led to several conclusions. These are laid out to address each previously defined research section (i.e., external communication impact, light bar factors, and testing scenarios).

External Communication Impact

Participants' behavior did not change once they understood what the external communication displays were trying to convey.

It took participants more than 12 exposures to the light bars to understand what the pattern meant. After 16 exposures, the largest percentage of participants (25/40, 63%) understood all three patterns. Only a little over half of the participants could correctly interpret the meanings throughout the session. Participants stated that they had difficulty dividing their attention between two L4 AVs at an intersection and between multiple vehicles with different light bars. 'Yielding' and 'Driver' patterns were the first two patterns for participants to interpret. Some participants *only* noticed the Yield and Drive pattern (15/40, 38%) across the entire study. The 'Ready' pattern was the most difficult and last for participants to understand. The pattern length was too short for participants to catch.

Participants found it challenging to focus on a vehicle's light bars when multiple vehicles were fighting for their attention in the same crossing vicinity. As a result, participants often chose one vehicle in their immediate path to focus on. As a result, participants gave up looking at the vehicle that was not directly impacting their intended crossing path and prioritized their attention on the most relevant vehicle. However, when participants did look back and forth, they often missed the vehicle's light bar patterns.

Participants expressed concern when both vehicles crossed their intended path and they had to interpret two light bars. The light bar patterns were often not synchronized (because they were completing different movements), and the various viewing sequences were confused. Thus they would miss the light patterns on the vehicles because they would turn their head back and forth to look at the other vehicles. During this movement, they missed the short Ready pattern change. This further suggests that simplifying the communication displays and improving visibility will be critical moving forward.

Light Bar Factors

Participants' crossing decision was not impacted by the condition (color or location). While the conditions did impact perceived system desirability, there was no significant difference among conditions regarding crossing decisions and behaviors.

However, in the light bar desirability findings, the color preference was nearly split between teal and white, with a few more participants preferring teal. The windshield light bar was preferred over the grill location due to greater ease of visibility. The windshield location was more visible on top of the vehicle because that is where participants looked for human interaction from the driver, and no other distracting lights were taking their attention. Less than half of the participants did not prefer the windshield location. Some stated the windshield location was more difficult to see or made the light bar seem too similar to police lights.

Some participants preferred the grill over the windshield location because the light bars were easier to see in this location. However, many participants did not prefer the grill location. Participants said the grill location was too low, and their eyes were not generally drawn to the grill for information. Additionally, participants were not looking at the grill because they usually looked at the windshield to find the driver. Furthermore, the light bar looked like a decoration in this location, and participants could not differentiate the light bar from an accent light. Finally, participants stated the surrounding area had too much happening already. Specifically, the surrounding chrome and headlights made it hard to see.

Participants did not prefer light bars that blended into the vehicle because they were difficult to discern from the vehicle. Based on different light bar conditions, participants did not prefer the light bars in the grill because they blended in with the headlights. They also did not prefer the thin windshield light bar because it blended in with the vehicle's roofline. Instead, participants preferred the thick light bars because they were more visible. These light bars were more visible because they were prominent from the vehicle and easily got participants' attention. It is important to note that a few participants stated that they would want the lights to stand out from the vehicle when these light bars are first introduced into the roadway network. Then over time, once pedestrians and drivers learned to recognize them, they could become smaller and more integrated into vehicles. However, a few participants preferred the lights that blended in with the vehicle because that was more aesthetically pleasing.

Testing Scenarios

A portion of the population was willing to cross in front of an L4 AV while it was still slowing to a stop; another portion of the population refused to cross in front of the L4 AV in all scenarios. These two extreme population groups need to be considered when planning future external communication designs, potentially by simplifying the displays and using preferred display locations.

Limitations of the Study

There are several limitations to this study. Participants had no prior knowledge of the lights, their location, pattern, or color. Participants were also never taught the meaning of the different light patterns throughout the entirety of the session. A baseline condition was used where one of the L4/5 AVs did not have its light bars turned on. However, there was no human driver baseline condition included.

The triggering of the light patterns was manually configured by the drivers concealed in the seat suit costume. Since the lights were manually triggered, consistency across trials may have varied. For future studies, the patterns should be wired to trigger when the human operator places their foot on the brake and be programmed to the speedometer. When the vehicle reaches zero, the corresponding pattern would be triggered.

The study was conducted during the daylight hours, so light bars were not viewed under nighttime conditions. Additionally, due to safety concerns, the study was not performed under precipitation conditions. Cloud coverage may also affect visibility of the lights.

Additional Products

The Education and Workforce Development (EWD) and Technology Transfer (T2) products created as part of this project can be downloaded from the project page on the Safe-D website. The final project dataset is located on the Safe-D Collection of the VTTI Dataverse.

Education and Workforce Development Products

This project provided students with the opportunities to take part in high-fidelity vehicle research throughout all phases of the experiment. Both undergraduate and graduate students were heavily involved in the performing the literature review, developing the research plan, conducting research, analyzing the data, and final report delivery. Throughout the entire process, students took on primary responsibility of the project and ensured it adhered to VTTI safety policies.

Students also gained vital public speaking skills through presenting the research plan, analysis updates, and final deliverable presentation to all key stakeholders (State Farm, Daimler, and Ford), and even an international organization (ISO). Additionally, students expanded their technical writing experience through a paper submission to *Transportation Research Record* (TRR) 2023. Visual external communication signals are being heavily investigated in the U.S. as well as abroad. The research team is involved in domestic and international committees exploring the potential for standardization in the design and deployment of L4/5 AV external communication.

Technology Transfer Products

This project produced an abstract which was submitted and accepted for presentation at the Transportation Research Board's annual meeting (TRB 2022 and TRB 2023). The project was presented at the Virginia Tech Industrial and System Engineering departments annual research symposium in April 2021 by the undergraduate researchers. Additionally, an academic journal article will be submitted to TRR 2023.

Data Products

A subset of data collected as part of the study is available via the Safe-D collection on the VTTI Dataverse. This data includes crossing decision, glance data for both L4/5 AVs, distance of crossing decision, vehicle condition (i.e., light bar color, light bar location, light bar thickness), number of decisions, and number of vehicles in participants' intended crossing path.

References

Basantis, A. (2019). *Assessing Alternate Approaches for Conveying Automated Vehicle 'Intentions.' November*. https://doi.org/10.13140/RG.2.2.28125.36324

Faas, S. M., Mathis, L. A., & Baumann, M. (2020). External HMI for self-driving vehicles: Which information shall be displayed? *Transportation Research Part F: Traffic Psychology and Behaviour*, *68*, 171–186. https://doi.org/10.1016/j.trf.2019.12.009

Habibovic, A., Lundgren, V. M., Andersson, J., Klingegård, M., Lagström, T., Sirkka, A., Fagerlönn, J., Edgren, C., Fredriksson, R., Krupenia, S., Saluäär, D., & Larsson, P. (2018). Communicating Intent of Automated Vehicles to Pedestrians. *Frontiers in Psychology*, *9*(August). https://doi.org/10.3389/fpsyg.2018.01336

Lee, Y. M., Madigan, R., Garcia, J., Tomlinson, A., Solernou, A., Romano, R., Markkula, G., Merat, N., & Uttley, J. (2019). Understanding the messages conveyed by automated vehicles. *Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019*, 134–143. https://doi.org/10.1145/3342197.3344546

SAE International. (2021). *SAE J3016 Levels of Driving Automation™ Refined for Clarity and International Audience*.

Schieben, A., Wilbrink, M., Kettwich, C., Madigan, R., Louw, T., & Merat, N. (2019). Designing the interaction of automated vehicles with other traffic participants: design considerations based on human needs and expectations. *Cognition, Technology and Work*, *21*(1), 69–85. https://doi.org/10.1007/s10111-018-0521-z

Appendix A. Testing Scenario Details

Scenario 1: 1 Vehicle Straight Crossing Path (1 Veh SCP)

Participants acted as a pedestrian. Vehicles started at 25 mph, and then at the 100-foot markings, they reduced speed to 10 mph. The AV-B arrived first, followed by the AV-A. The vehicles were perpendicular to each other. Vehicles stopped at the stop sign for 5-seconds. The AV-B proceeded straight first, followed by the AV-A. Across all trials, AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

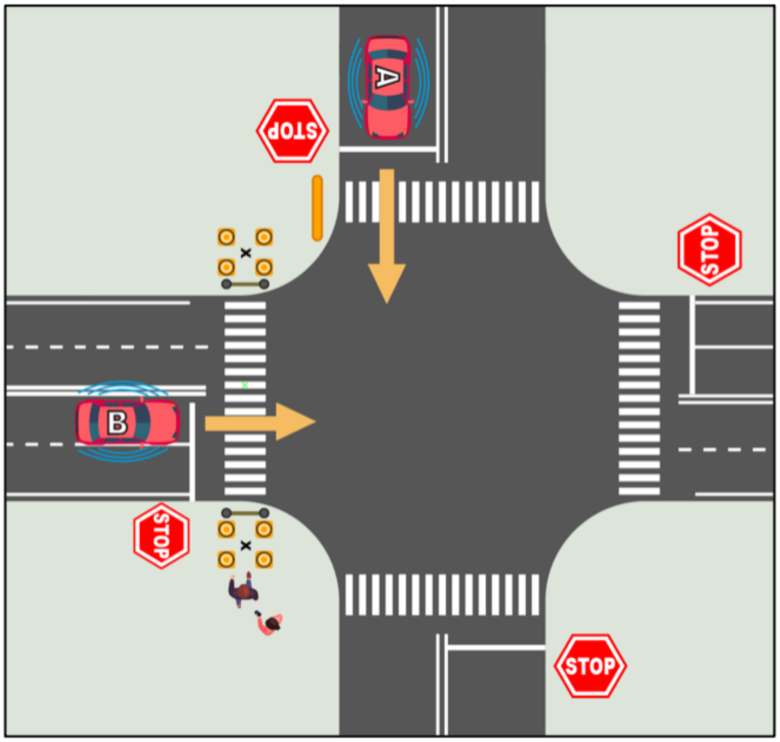

Figure 14. Scenario 1 aerial depiction.

Scenario 2: Right Turn (RT)

Participants acted as a pedestrian. Vehicles started at 25 mph, and then at the 100-foot markings, slowed to 10 mph. The AV-B arrived first, followed by the AV-A. Vehicles stopped at the stop sign for 5 seconds. The AV-B proceeded straight first, followed by the AV-A. Across all trials, AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

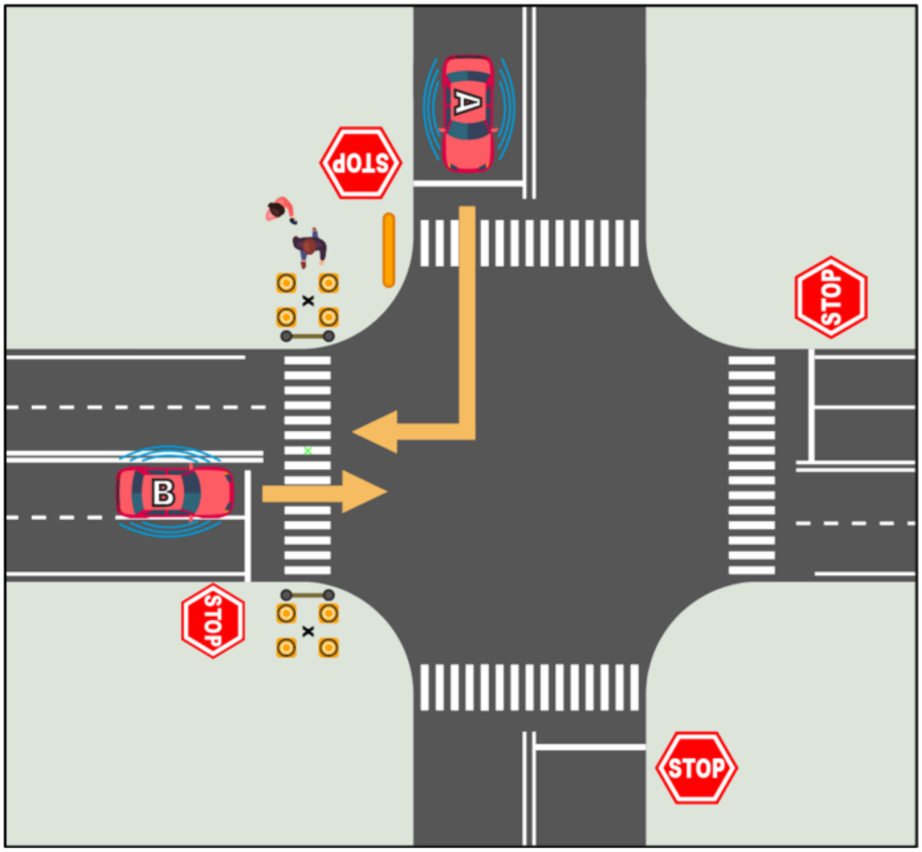

Figure 15. Scenario 2 aerial depiction.

Scenario 3: 2 Vehicle Straight Crossing Path (2 Veh SCP)

Participants acted as a pedestrian. Vehicles started at 25 mph, and then at the 100-foot markings, slowed to 10 mph. The AV-B arrived first, followed by the AV-A. Vehicles were facing each other in the intersection. Vehicles stopped at the stop sign for 5 seconds. The AV-B proceeded straight first, followed by the AV-A. Across all trials, AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

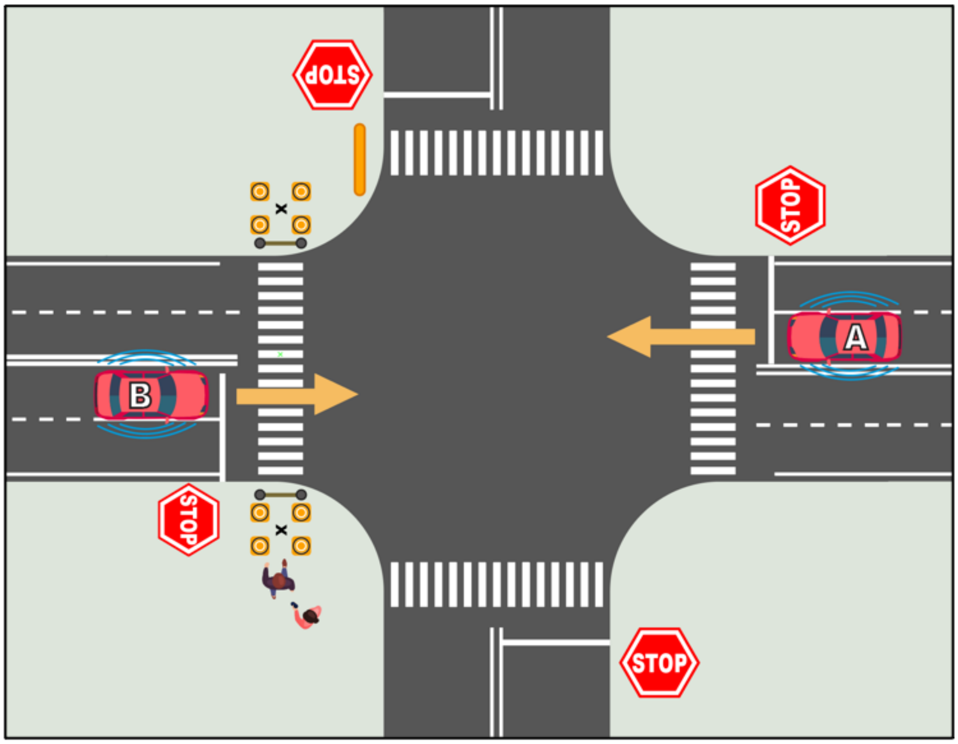

Figure 16. Scenario 3 aerial depiction.

Scenario 4: Lane Change and Simulated Pedestrian

Participants were in control of the simulated pedestrian. Participants indicated when would be the last moment they felt the simulated pedestrian could safely make it across the street. Vehicles proceed down the road at 35 mph. AV-A drove down the street with the AV-B behind. At a designated 100-foot marker from the pedestrian, the AV-B switched to the left lane and passed the AV-A. The AV-A (in the right lane) "detected" the simulated pedestrian crossing the street and would come to a controlled stop. The AV-B (which passed the AV-A in the left lane) "detected" the simulated pedestrian crossing the street and yielded with a hard brake (.7 g brake) to maintain distance. Across all trials, AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

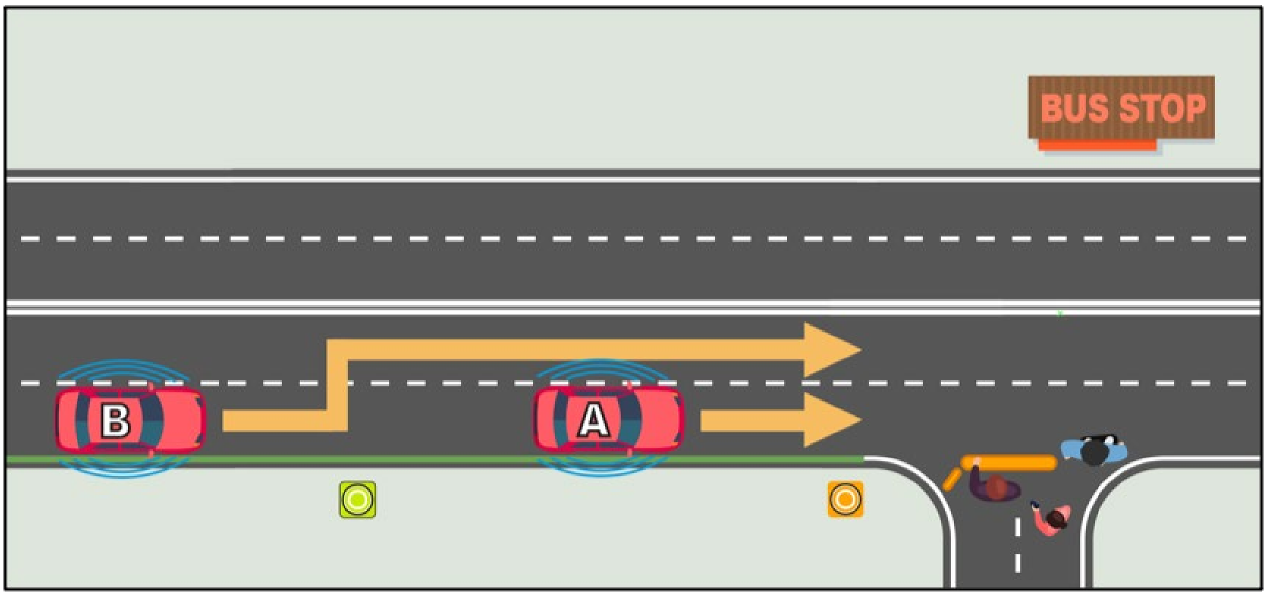

Figure 17. Scenario 4 aerial depiction.

Scenario 5: Lane Change with Expert Pedestrian

Participants acted as a passenger in the front passenger seat of a testing vehicle. The moderator operated this testing vehicle. The participant vehicle was on the other side of the intersection facing the L4/5 AV. Vehicles started at 25 mph, and then at the 100-foot markings, slowed to 10 mph. AV-A drove down the street with AV-B behind. At a designated 100-foot marker from the pedestrian, AV-B switched to the left lane and passed AV-A. AV-A arrived at the intersection first, followed by AV-B. At the intersection, there was an "expert pedestrian" waiting to cross the street. Once both vehicles emulating SAE L4/5 AV displays came to a complete stop, the "expert pedestrian" crossed the street. The “expert pedestrian” was a trained VTTI employee. The participants watched this interaction from the opposite side of the intersection. Across all trials, AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

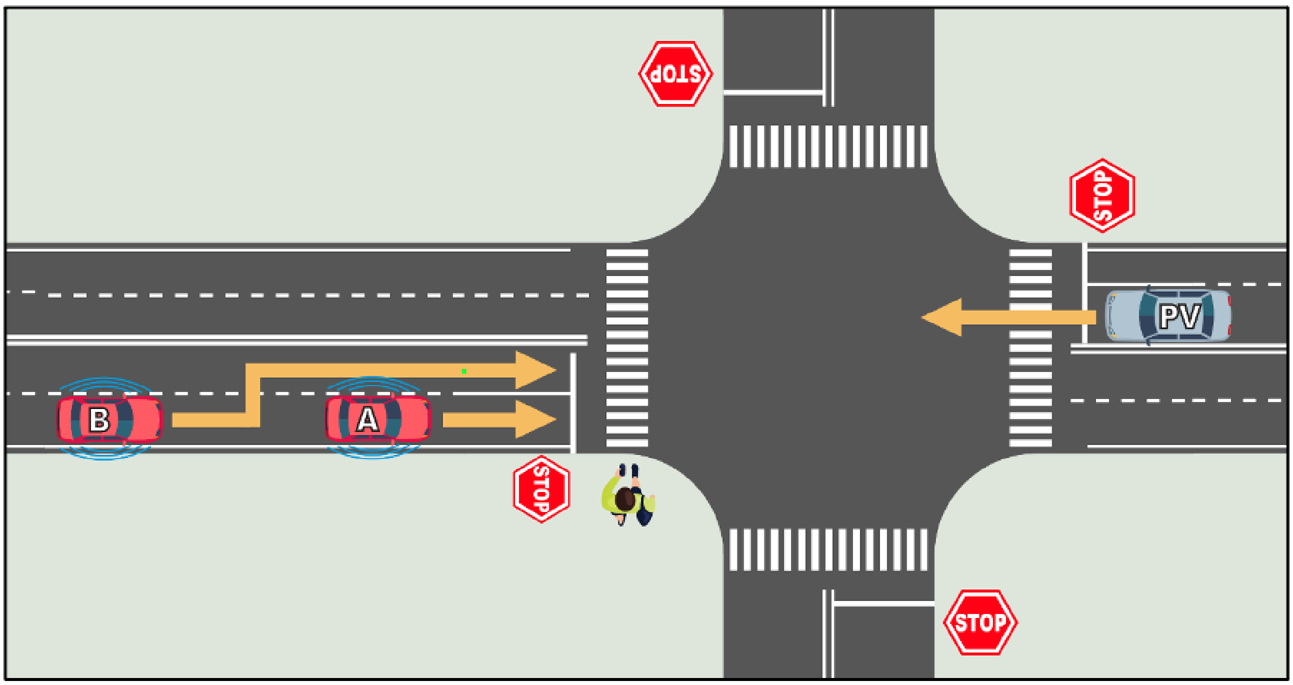

Figure 18. Scenario 5 aerial depiction.

Scenario 6: 4-way Stop and Vehicle Left Turn

Participants acted as a passenger in the front passenger seat of a testing vehicle. The moderator operated this testing vehicle. The participant vehicle was perpendicular to the L4/5 AVs. Vehicles started at 25 mph, and then at the 100-foot markings, slowed to 10 mph. When the L4/5 AVs reached the intersection, they stopped for 5 seconds. AV-B proceeded straight through the intersection, and then AV-A made a left-hand turn. The participants watched this interaction from the intersection. Across all trials, AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

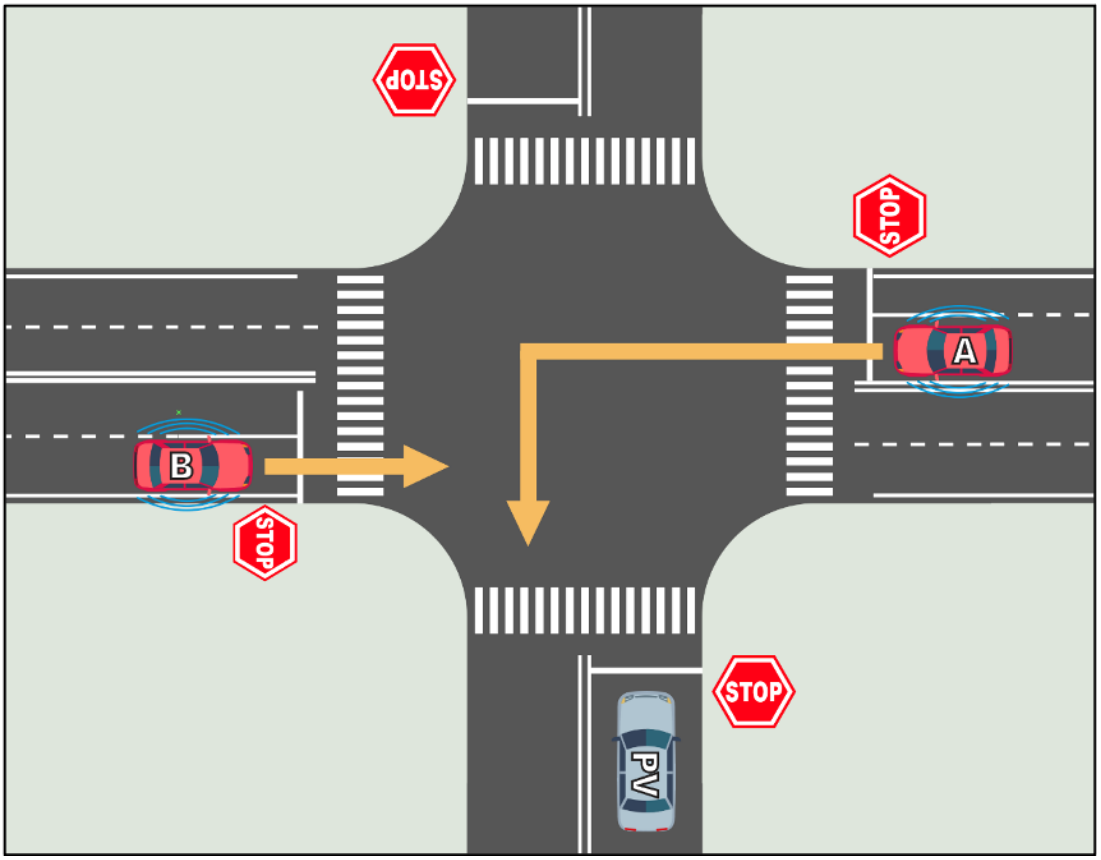

Figure 19. Scenario 6 aerial depiction.

Scenario 7: Construction Zone with Active Flagger

Participants acted as a passenger in the front passenger seat of a testing vehicle. The moderator operated this testing vehicle. The participant vehicle was on the other side of the construction zone facing the L4/5 AVs. Vehicles started at 25 mph, and then at the 100-foot markings, slowed to 10 mph. The participant vehicle drove to the construction zone and was instructed to wait by the flagger. The flagger held a stop sign and waved a flag to indicate to the driver of the participant vehicle to stop. The flagger held the participant vehicle stopped while the scenario with AV-A and B occurred. While the participant's vehicle was stopped, the participant was instructed to watch the L4/5 AVs. The L4/5 AVs approached the intersection and were perpendicular. AV-A made a right-hand turn and was stuck in a right lane that ended and merged into a single lane. At this point, AV-B proceeded straight through the intersection and straight through the construction zone. After this pinch point and AV-B passed, AV-A merged into the single lane and proceeded straight. Across all trials, AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

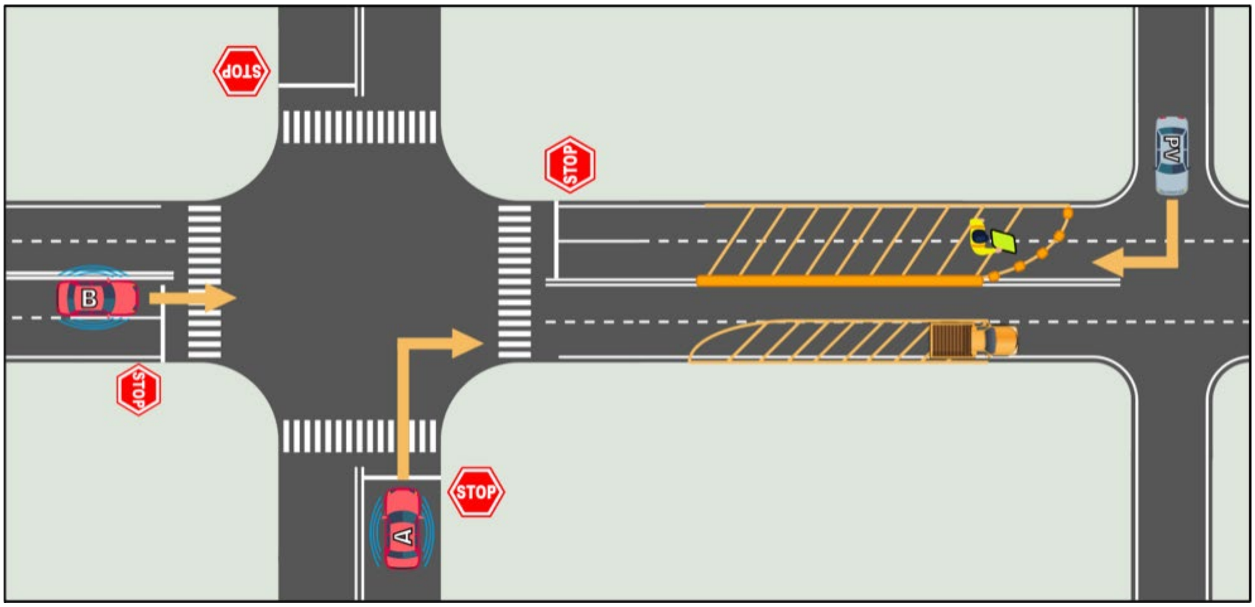

Figure 20. Scenario 7 aerial depiction.

Scenario 8: Mid-block Lane Change

Participants acted as a passenger in the front passenger seat of a testing vehicle. The moderator operated this testing vehicle. This is the only scenario where the participant vehicle was driving alongside the L4/5 AVs. All vehicles started in the cul-de-sac, where the AV-A was first, followed by the participant vehicle and the AV-B behind in the right lane. In the first segment, the AV-A yielded at the 200-foot marker from the intersection. The AV-B merged into the left lane and passed both the AV-A and the participant vehicle. After the AV-A yielded, it began to drive down the road towards the intersection. Both L4/5 AVs were driving side-by-side, and the participant vehicle was still behind the AV-A. The L4/5 AVs arrived at the intersection simultaneously, stopped for five seconds, and then proceeded through.

There was a solid LED strip mounted on the back of the AV-A. The LED light bar did not have a pattern and remained static the entire driving scenario. The viewing interaction of the participant was only possible (more or less) through viewing the side mirrors and the rear-view mirror. Each participant adjusted the mirrors before the start of each trial. AV-B was always behind the participant. AV-A and AV-B never switched locations due to the complexity of scenarios and to maintain participant and driver safety.

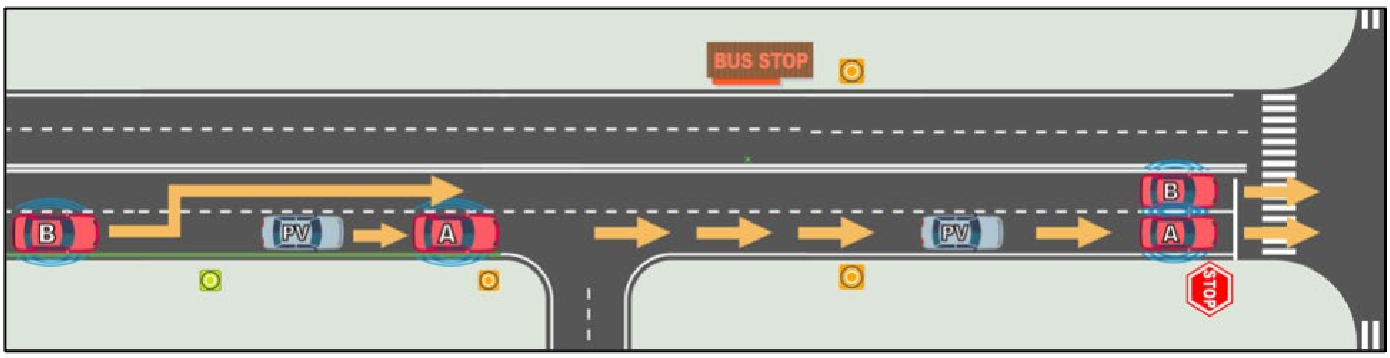

Figure 21. Scenario 8 aerial depiction.

Appendix B. Glance Data Significant Pairwise Comparison Results

Table 5. Glance Data Significant Pairwise Comparison Results

Trial	Trial Compared	Mean Difference	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
1	2	-0.226	0.244	1.000	-1.063	0.611
	3	-0.258	0.217	1.000	-1.003	0.487
	4	-0.484	0.289	1.000	-1.475	0.507
	5	1.065*	0.207	0.000	0.355	1.774
	6	1.000*	0.185	0.000	0.364	1.636
	7	-.774*	0.226	0.050	-1.548	-7.651E-05
	8	-0.581	0.235	0.548	-1.388	0.226
2	1	0.226	0.244	1.000	-0.611	1.063
	3	-0.032	0.234	1.000	-0.835	0.770
	4	-0.258	0.304	1.000	-1.300	0.784
	5	1.290*	0.162	0.000	0.735	1.845
	6	1.226*	0.152	0.000	0.706	1.746
	7	-0.548	0.249	0.994	-1.402	0.305
	8	-0.355	0.256	1.000	-1.234	0.524
3	1	0.258	0.217	1.000	-0.487	1.003
	2	0.032	0.234	1.000	-0.770	0.835
	4	-0.226	0.273	1.000	-1.162	0.711
	5	1.323*	0.163	0.000	0.763	1.882
	6	1.258*	0.173	0.000	0.664	1.852
	7	-0.516	0.217	0.669	-1.260	0.227
	8	-0.323	0.238	1.000	-1.139	0.494
4	1	0.484	0.289	1.000	-0.507	1.475
	2	0.258	0.304	1.000	-0.784	1.300
	3	0.226	0.273	1.000	-0.711	1.162
	5	1.548*	0.262	0.000	0.651	2.445
	6	1.484*	0.245	0.000	0.645	2.323
	7	-0.290	0.259	1.000	-1.178	0.598
	8	-0.097	0.309	1.000	-1.156	0.962
5	1	-1.065*	0.207	0.000	-1.774	-0.355
	2	-1.290*	0.162	0.000	-1.845	-0.735
	3	-1.323*	0.163	0.000	-1.882	-0.763
	4	-1.548*	0.262	0.000	-2.445	-0.651
	6	-0.065	0.079	1.000	-0.337	0.208
	7	-1.839*	0.174	0.000	-2.436	-1.242

Trial	Trial Compared	Mean Difference	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b] Lower Bound	95% Confidence Interval for Difference[b] Upper Bound
	8	-1.645*	0.194	0.000	-2.311	-0.979
6	1	-1.000*	0.185	0.000	-1.636	-0.364
	2	-1.226*	0.152	0.000	-1.746	-0.706
	3	-1.258*	0.173	0.000	-1.852	-0.664
	4	-1.484*	0.245	0.000	-2.323	-0.645
	5	0.065	0.079	1.000	-0.208	0.337
	7	-1.774*	0.172	0.000	-2.363	-1.186
	8	-1.581*	0.184	0.000	-2.212	-0.949
7	1	.774*	0.226	0.050	7.651E-05	1.548
	2	0.548	0.249	0.994	-0.305	1.402
	3	0.516	0.217	0.669	-0.227	1.260
	4	0.290	0.259	1.000	-0.598	1.178
	5	1.839*	0.174	0.000	1.242	2.436
	6	1.774*	0.172	0.000	1.186	2.363
	8	0.194	0.215	1.000	-0.542	0.929
8	1	0.581	0.235	0.548	-0.226	1.388
	2	0.355	0.256	1.000	-0.524	1.234
	3	0.323	0.238	1.000	-0.494	1.139
	4	0.097	0.309	1.000	-0.962	1.156
	5	1.645*	0.194	0.000	0.979	2.311
	6	1.581*	0.184	0.000	0.949	2.212
	7	-0.194	0.215	1.000	-0.929	0.542

Appendix C. Pairwise Comparison of Pattern Notice

Table 6. Pairwise Comparison of Pattern Notice

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Adj. Sig
1-12	-0.350	0.094	-3.705	0.032
1-16	-0.350	0.094	-3.705	0.032
17-10	0.550	0.094	5.822	0.000
17-11	0.375	0.094	3.969	0.011
17-12	0.675	0.094	7.145	0.000
17-13	0.625	0.094	6.615	0.000
17-14	0.600	0.094	6.351	0.000
17-15	0.600	0.094	6.351	0.000
17-16	0.675	0.094	7.145	0.000
17-2	0.450	0.094	4.763	0.000
17-3	0.550	0.094	5.822	0.000
17-4	0.550	0.094	5.822	0.000
17-5	0.475	0.094	5.028	0.000
17-6	0.625	0.094	6.615	0.000
17-7	0.650	0.094	6.880	0.000
17-8	0.650	0.094	6.880	0.000
18-10	0.500	0.094	5.292	0.000
18-12	0.625	0.094	6.615	0.000
18-13	0.575	0.094	6.086	0.000
18-14	0.550	0.094	5.822	0.000
18-15	0.550	0.094	5.822	0.000
18-16	0.625	0.094	6.615	0.000
18-2	0.400	0.094	4.234	0.004
18-3	0.500	0.094	5.292	0.000
18-4	0.500	0.094	5.292	0.000
18-5	0.425	0.094	4.499	0.001
18-6	0.575	0.094	6.086	0.000
18-7	0.600	0.094	6.351	0.000
18-8	0.600	0.094	6.351	0.000
9-10	-0.400	0.094	-4.234	0.004
9-12	-0.525	0.094	-5.557	0.000
9-13	-0.475	0.094	-5.028	0.000
9-14	-0.450	0.094	-4.763	0.000
9-15	-0.450	0.094	-4.763	0.000
9-16	-0.525	0.094	-5.557	0.000
9-3	0.400	0.094	4.234	0.004
9-4	0.400	0.094	4.234	0.004
9-6	0.475	0.094	5.028	0.000
9-7	0.500	0.094	5.292	0.000
9-8	0.500	0.094	5.292	0.000