Standardized Performance Evaluation of Vehicles with Automated Capabilities

May 2020 | Final Report

SAFE-D
SAFETY THROUGH DISRUPTION

Transportation Institute
Virginia Tech

Texas A&M Transportation Institute

San Diego State University
Leadership Starts Here

Disclaimer

TECHNICAL REPORT DOCUMENTATION PAGE

<table>
<tr><td>1. Report No.
VTTI-00-020</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Standardized Performance Evaluation of Vehicles with Automated Capabilities</td><td colspan="2">5. Report Date
December 2019</td></tr>
<tr><td colspan="2">6. Performing Organization Code:</td></tr>
<tr><td colspan="2">7. Author(s)
Alexis Basantis
Leslie Harwood
Zachary Doerzaph
Luke Neurauter</td><td colspan="2">8. Performing Organization Report No.
Report VTTI 00-020</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address:
Safe-D National UTC
Virginia Tech Transportation Institute
3500 Transportation Research Plaza
Blacksburg, VA 24060</td><td colspan="2">10. Work Unit No.</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
69A3551747115/Project VTTI-00-020</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
Office of the Secretary of Transportation (OST)
U.S. Department of Transportation (US DOT)</td><td colspan="2">13. Type of Report and Period
Final Research Report</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code</td></tr>
<tr><td colspan="4">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.</td></tr>
<tr><td colspan="4">16. Abstract
Advanced driver-assistance systems (ADAS) are becoming widely available in the new vehicle landscape, increasing of both vehicle occupants' and other road users' safety. In some vehicles, longitudinal and lateral positioning under certain conditions can be maintained, designating them as having either SAE level 1 (L1) or level 2 (L2) automated features. By developing a standardized set of tests to be applied to current L1 and L2 vehicles, while keeping the future advancement of automation in mind, these vehicles' system performance, feature limitations, and performance consistency can be systematically evaluated. This project sought to develop an easily implementable, standardized set of testing procedures that could be quickly and inexpensively performed on automated vehicles to characterize their feature capabilities and limitations. Such information is useful to private or public organizations interested in a standardized approach to classifying vehicle capabilities, whether for informing the expectation of operators, or for cataloging and learning from the variety of implementation alternatives. Although not the primary purpose, this project may also help inform efforts to develop certification or other standardized vehicle performance efforts. The results of this project showed that specific roadway factors affected automated feature performance and that there was significant performance variability across test vehicles.</td></tr>
<tr><td colspan="2">17. Key Words
Automated, performance, evaluation, standardized, testing, capabilities</td><td colspan="2">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.</td></tr>
<tr><td>19. Security Classif. (of this report)
Unclassified</td><td>20. Security Classif. (of this page) Unclassified</td><td>21. No. of Pages
[insert # pages]</td><td>22. Price
$0</td></tr>
</table>

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

Abstract

Advanced driver-assistance systems (ADAS) are becoming widely available in the new vehicle landscape, increasing of both vehicle occupants' and other road users' safety. In some vehicles, longitudinal and lateral positioning under certain conditions can be maintained, designating them as having either SAE level 1 (L1) or level 2 (L2) automated features. By developing a standardized set of tests to be applied to current L1 and L2 vehicles, while keeping the future advancement of automation in mind, these vehicles' system performance, feature limitations, and performance consistency can be systematically evaluated. This project sought to develop an easily implementable, standardized set of testing procedures that could be quickly and inexpensively performed on automated vehicles to characterize their feature capabilities and limitations. Such information is useful to private or public organizations interested in a standardized approach to classifying vehicle capabilities, whether for informing the expectation of operators, or for cataloging and learning from the variety of implementation alternatives. Although not the primary purpose, this project may also help inform efforts to develop certification or other standardized vehicle performance efforts. The results of this project showed that specific roadway factors affected automated feature performance and that there was significant performance variability across test vehicles.

Acknowledgements

The authors would like to thank Craig Fedor, Eric Glenn, and Jacob Walters for their assistance in the development and execution of the project. In addition, we would like to thank all of the different teams at VTTI for their time and efforts.

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

Automated vehicle technologies are evolving at a rapid pace and implementations of various automated features vary widely across different vehicle makes and models. The Virginia Tech Transportation Institute (VTTI) has worked for years with private and public partners to test Advanced Driver Assistance Systems (ADAS)-equipped vehicles with automated capabilities and is now seeing a rapid increase in the prevalence of such features. As expected, the capabilities of ADAS-equipped vehicles arriving at VTTI for further research and testing vary broadly, with the possibility of each handling roadway environments and scenarios differently. For example, one vehicle may only alert the human driver to an oncoming threat, expecting the driver to take control of the vehicle, while another vehicle may employ automatic braking to mitigate the same oncoming threat. In the near future, vehicles may attempt evasive swerves to avoid such a threat. Organizations using many different vehicles and those involved in standardization or design will likely benefit from a systematic approach to exploring implementation and performance differences and cataloging the variety of systems available in the marketplace, including an inventory of their differences.

The near-term goal of this project was to develop an easily implementable, standardized set of testing procedures that could be quickly and inexpensively performed on automated vehicles new to public and private research institutions to determine their feature capabilities and limitations. In addition, researchers were curious whether these simple tests could create an adequate enough baseline metric to eventually build a refined set of these tests that could be performed in conjunction with a more robust certification process to properly evaluate performance capabilities across of all levels of vehicle automation. In addition, researchers wanted to better understand current capabilities of automated vehicle features and the variations between features developed by different OEMs.

To develop these procedures, the research team conducted a review of available literature to gain an understanding of current standardized vehicle tests and instances where automated features may fail [1, 2, 3]. In addition, focus groups were held at VTTI, where the preliminary set of tests were presented to automated vehicle industry experts to leverage their opinions and experiences. Questions asked during the focus group can be found in Appendix A: Focus Group Questions. The findings of the focus groups were used to refine the preliminary test designs, and the resultant tests were then applied to a fleet of SAE International level 1 and level 2 (L1 and L2) vehicles using both closed test-tracks (e.g., the Virginia Smart Roads at VTTI) and public roads. A portfolio was created outlining the performance of each vehicle, as well as its performance relative to other vehicles within the fleet. The aim is for this portfolio to be a living document that will include more vehicles as automated systems evolve and new vehicles become available for testing.

The goal of this project was to answer the following research questions:

1. What testing should be conducted to evaluate both current and forthcoming capabilities of vehicles with automated ADAS features?
2. How do currently available vehicles with automated driving features perform under the proposed standardized set of evaluations?

Background

Automated vehicles are becoming more prevalent as consumers recognize the potential they offer for increased convenience and safety, and manufacturers deploy systems intended to meet consumer demands. However, not all automated vehicles are the same. There are different levels of automated vehicle capabilities, ranging from vehicles providing minimal driver assistance to those providing full automation. SAE International has developed a scale that categorizes these varying levels of automation within vehicles, as shown in Figure 1 [4].

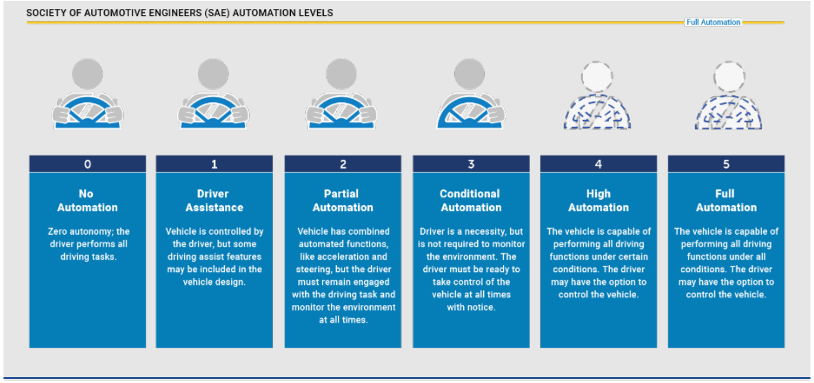


Figure 1. SAE International levels of automation [5].

Higher levels of automated vehicles are on the horizon; however, the most widely available automated vehicles currently on the consumer-market are designated as level 1 (L1) or level 2 (L2) vehicles. These vehicles have the ability to control the vehicle's lateral (e.g., lane keeping assist [LKA]) and longitudinal (e.g., adaptive cruise control [ACC], automatic emergency braking [AEB]) dynamics, without any input from the user. However, users are still required to be constantly engaged in the driving task, as takeover may be necessary at any point in time. Vehicles with automated features have the potential to significantly reduce or mitigate vehicle crashes and have been proven to increase the safety of vehicle occupants and other roadway users [6].

Although these features have been proven to increase safety, minimal previous research examining these features' variability and limitations has been conducted. Prior to the beginning of this study, most research done in this area focused on extreme edge-case scenarios or theoretical modeling of automated vehicles limited to the most common accident scenarios [7, 8, 9]. No real-world experimentation had been conducted to better understand how currently-available automated features perform in daily driving scenarios. Since this project began, several other high-profile organizations, such as the Insurance Institute for Highway Safety and the European New Car

Assessment Programme, have also begun developing standardized tests for automated features [10, 11].

Methods

Study Vehicles

Developed tests were applied to a fleet of the most advanced automated vehicles available on the consumer market in 2018. A summary table outlining the study vehicles used and their automated technology packages can be found in Table 1. Automated feature nomenclature for each vehicle can be found in Appendix B: Fleet Vehicles and Automated Driver Assistance System (ADAS) Packages.

Table 1. Vehicles Used in the Study

Vehicle Make/Model	Automated Feature Package Included in Vehicle
2017 Audi Q7 Premium Plus 3.0 TFSI Quattro	Driver Assistance Package
2015 Infiniti Q50 3.7 AWD Premium	Technology, Navigation, and Deluxe Touring Package
2016 Mercedes-Benz E350 Sedan	Premium Package, Driver Assistance Package
2015 Tesla Model S P90D AWD	Autopilot Convenience
2016 Volvo XC90 T6 AWD R-Design	Convenience Package
2018 Cadillac CT6 AWD 3.6L Engine Premium Luxury	Driver Awareness and Convenience Package, Super Cruise Package
2018 Tesla Model X	Performance Package

As mentioned previously, the set of tests were designed considering current academic literature and in collaboration with a panel of automated and connected vehicle experts at VTTI. In addition, due to the higher prevalence of ACC, LKA, and AEB in the current market, the tests applied in this study primarily focused on these features. All tests used a full-factorial design approach to iterate through different scenario parameters.

Tests were conducted on the Virginia Smart Roads located at VTTI. The Virginia Smart Roads are a collection of controlled-access test beds simulating a variety of roadway environments and were designed to meet Virginia Department of Transportation (VDOT) standards. For this study, both the Highway, which allows for higher driving speeds and multiple turn-arounds, and the Surface Street, which provides lower speed roadway features, sections of the road were used.

Trained experimenters documented each test trial, recording vehicle behavior, test parameters, and general observations regarding vehicle performance. Performance variables such as speed, GPS coordinates, and acceleration were collected using a VTTI-developed data acquisition system installed in the vehicles for an unrelated effort. However, as our goal for this project was to enable a quick and inexpensive evaluation of ADAS capability, we focused on the development of a robust observer protocol rather than detailed analysis of the parametric data.

The testing included dynamic vehicle maneuvers at typical speeds for each maneuver. For this reason, prior to experimentation, testing and safety procedures underwent an extensive review from VTTI's Research Review Committee. In addition, safety protocols, expected vehicle response to new stimuli (Appendix C: Expected Vehicle Performance), "what-if" events, and bailout procedures were also developed for each test. During testing, all vehicle maneuvers were performed by highly experienced drivers who routinely drive new production and test vehicles at VTTI. Researchers kept constant communication with each other throughout tests via two-way radios. On-road researchers, external to the vehicle, remained a safe distance away from test vehicles while they were in motion.

Attempts were also made to reduce variability between the ADAS configurations across the vehicles. For tested vehicles, configuration settings included ACC feature availability and headway settings. All features were either turned on or experimentally controlled as appropriate for the test being performed. There were 3–7 different vehicle headway setting options (time based) in the test fleet. Through vehicle manuals, research, and physical testing, each headway option was determined, and three variations, which corresponded to long (headway of ~3 s), medium (headway of ~2 s), and short (headway of ~1 s), were identified for use during testing.

ACC and LKA Test Scenarios

The Highway section of the Virginia Smart Roads was used to test ACC and LKA vehicle features. Subsequent sections describe these Highway tests in more detail. Additional descriptive scenario diagrams and images can be found in Appendix D: Additional Test Scenario Images. Test parameters for each scenario can be found in Appendix E: Test Parameters.

ACC Curve Test

The ACC Curve test was performed to evaluate the effectiveness of ACC and LKA features for each vehicle. In this test, the test vehicle followed a lead vehicle traveling at a constant speed (20 mph, 25 mph, 30 mph) around curves on the Highway and Surface Street sections of the Virginia Smart Roads (example depicted in Figure 2), while ACC was set to a prescribed speed (15 mph, 20 mph, 25 mph) and following distance (medium, long). This route included curves with radii of 108 ft. (32.9 m) and 205 ft. (62.5 m) on the Highway Section and 295 ft. (89.9 m) and 301 ft. (91.7 m) on the Surface Street section of the Smart Roads. Curve radii was defined as the distance from the center of the curve to the lane's outmost edge.

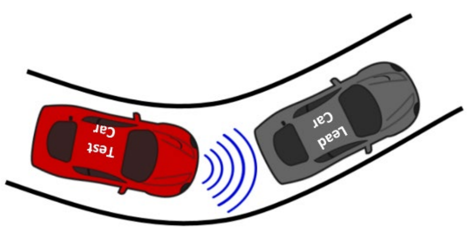


Figure 2. Diagram of ACC curve scenario around curve on the Highway section.

The speeds of the lead and test vehicle were paired and tests were performed sequentially (i.e., lead vehicle would travel at 15 mph while test vehicle traveled at 20 mph). Individual headway settings were applied to each paired speed variation (e.g., 15 mph/20 mph were conducted at both long and medium headway). Researchers evaluated whether the test vehicle tracked the lead vehicle throughout the entire curve either by cues on the visual human-machine interface (HMI) or increased test vehicle acceleration, which indicated a loss of tracking. If the test vehicle ceased tracking the lead vehicle at any point in the curve, the test vehicle driver was required to intervene and resume control of the driving task.

ACC Cut-in

The ACC Cut-in test was used to evaluate the test vehicle's ACC response to a quickly changing lead vehicle and the presence of multiple lead vehicles. This test was conducted on a straight mile the Highway section. During experimentation, the test vehicle, set to a constant ACC speed (35 mph, 45 mph, 55 mph) and following distance (medium, long), trailed a lead vehicle while a second lead vehicle (e.g., the "cut-in vehicle") traveled in the same direction but in the adjacent lane, as seen in Figure 3. Once both the test vehicle and primary lead vehicle reached steady state (i.e., both were holding the speed specified for that trial), the cut-in vehicle began to merge into the traveling lane between the test vehicle and the lead vehicle, as if the cut-in vehicle driver was changing lane without looking into their blind spot. The cut-in vehicle then held a constant speed of 30 mph, 40 mph, or 50 mph while remaining centered on the lane line, allowing experimenters in the test vehicle to examine the test vehicle's response to the new cut-in vehicle. Similar to the ACC curve test, the vehicle speeds were paired and both headway options were applied to each speed variation.

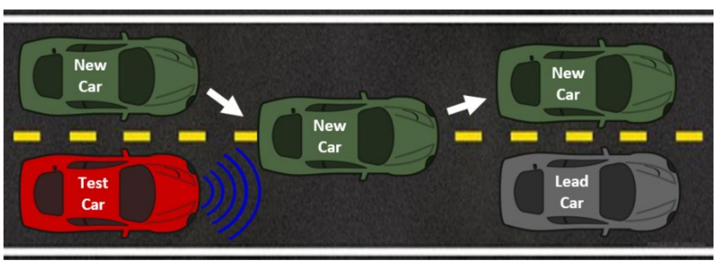


Figure 3. ACC Cut-In scenario diagram.

ACC Cut-out

Similar to the ACC Cut-in test, the ACC Cut-out test was used to further evaluate the test vehicle's response to a quickly changing lead vehicle, that, in some cases, was traveling at a drastically different speed. The lead vehicle (e.g., the "cut-out" vehicle) traveling at 20 mph, 30 mph, or 40 mph was followed by the test vehicle set to a specified ACC traveling speed (25 mph, 35 mph, 45 mph) and headway setting (medium, long). About a half-mile away from the test and lead vehicles, another vehicle (e.g., the "revealed" vehicle or "slow car" as seen in Figure 4) was traveling at a specified speed of 15 mph, 10 mph, or 0 mph. The test vehicle and cut-out vehicle both traveled toward the revealed vehicle, and just as the pair approached, the lead vehicle changed lanes, revealing the slower moving vehicle (i.e., "slow car"), as seen in Figure 4. Researchers made note of how the test vehicle's ACC features reacted to the presence of a new lead vehicle with a significantly lower speed. Each revealed vehicle speed and headway setting was applied to each test vehicle and lead vehicle speed pairing.

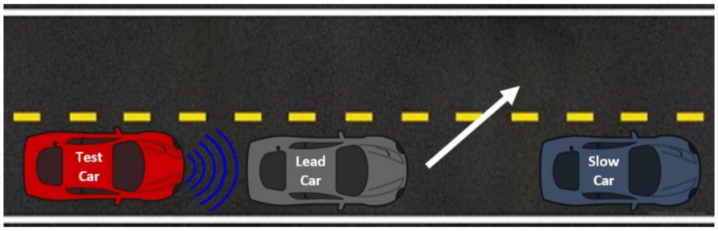


Figure 4. ACC Cut-Out scenario diagram.

Since this test involved higher speed differences, it posed a greater risk to researchers and vehicles. To increase safety, all trials were performed on an uphill portion of the Highway section of the Smart Road. Also, to ensure safety, in trials when the revealed vehicle was stationary (i.e., 0 mph), no researchers were present in the revealed vehicle and trials were only performed at approach speeds of 20 mph and 30 mph, decreasing the maximum speed delta.

Stop and Go

The Stop and Go test was used to examine the test vehicle's ability to react to a lead vehicle constantly changing speeds, including decelerating to a stop and then starting again, similar to what is experienced in traffic. The test vehicle was positioned behind two lead vehicles and the ACC was set to a constant speed of 35 mph, as seen in Figure 5. At the start of the test, the two lead vehicles accelerated to 30 mph while a 2-s headway was maintained between the two. Once the lead vehicles reached a steady state, the in-vehicle researcher cued the lead vehicle drivers to decelerate to 15 mph. Once all vehicles decelerated and again reached a steady state, the lead vehicle drivers were directed to accelerate to 25 mph until a steady state was again achieved. Finally, the lead vehicle drivers were directed to decelerate to 0 mph. When all vehicles reached a complete stop, the lead vehicle drivers accelerated back to 30 mph. After the stop, the test vehicle's ACC was re-engaged so it could follow the lead vehicles as they accelerated back up to speed.

During testing, researchers noted the effectiveness of the test vehicle's ACC in matching the changing speeds and adjusting following distances. As the speeds were consistent across test iterations, headway setting was the only parameter altered across trials.

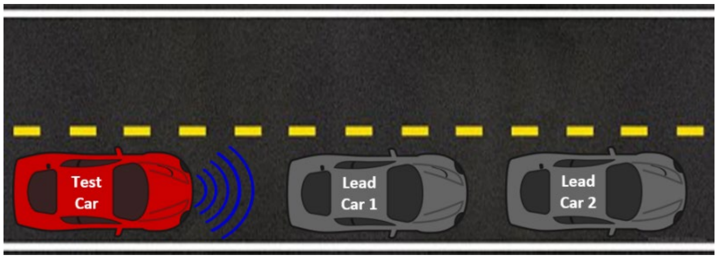


Figure 5. Configuration of vehicles in the stop and go test.

LKA Inattentiveness Test

The LKA Inattentiveness test evaluated the test vehicle's response to a driver taking their hands off of the steering wheel while ACC and LKA were engaged. The aim of this test was to determine how long a vehicle with active automated features would allow a driver to have their hands off the steering wheel and what types of strategies were employed to re-engage a driver in the driving task.

Test vehicle drivers set the ACC to the prescribed trial speed (45 mph, 55 mph) and headway setting (medium, long). At specified points along the road, marked with yellow traffic cones, the test vehicle driver removed their hands from the steering wheel and hovered them over the wheel as a safety precaution. Researchers recorded how long it took the vehicle to react to the lack of driver input, what type of warnings were presented (i.e., visual, auditory, haptic), and how long it took for the automated systems to disengage. If the vehicle features disengaged due to the lack of driver input, the test driver allowed the vehicle to drift out of the lane by one half of a car width (e.g., centered over the median). By allowing the vehicle to drift out of the lane, researchers were able to determine if the vehicle would take any sort of preventative measures to resist the lane drift, as shown in Figure 6.

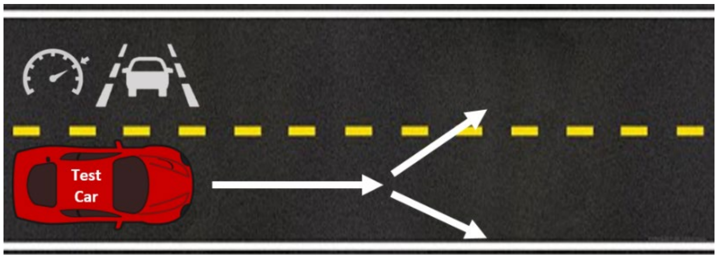


Figure 6. LKA Inattentiveness scenario diagram.

Low Speed ACC and AEB Test Scenarios

The Surface Street section of the Virginia Smart Roads was used to test ACC and AEB test vehicle fleet features. Additional descriptive scenario diagrams and images can be found in Appendix D: Additional Test Scenario Images. Test parameters for each scenario can be found in Appendix E: Test Parameters.

AEB Obstacle

The aim of this test was to determine each test vehicle's AEB capabilities when presented with a variety of objects directly in the driving path. Obstacles used consisted of both static (large pedestrian cut-out, small pedestrian cut-out, foam car) and dynamic (large pedestrian cut-out, small pedestrian cut-out) obstacles. The dynamic obstacles were controlled by an additional on-road researcher using VTTI's heavy vehicle remote evasive maneuvering device (Figure 7). As depicted in Figure 7, expert drivers drove the test vehicle at a speed of 25 mph, 35 mph, or 45 mph toward an obstacle set up directly in the driving path, or that was moved into the center of the traveling lane by an on-road researcher, as the test vehicle approached. When the test vehicle reached the obstacle, the driver was prepared to perform evasive maneuvers in case the system did not recognize the object as a hazard and brake accordingly.

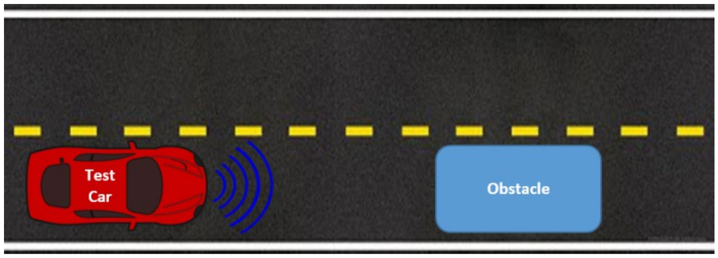


Figure 7. AEB Obstacle scenario diagram.

Lane Obstruction

Similar to the AEB obstacle test, the Lane Obstruction test was also used to determine the test vehicle's ability to recognize stationary obstacles in its path and to gauge its subsequent response. However, in this case, the vehicle's ACC features were engaged. In addition, the obstacles present were only halfway in the lane, as seen in Figure 8, rather than fully centered. Positioning the obstacle only halfway in the lane simulated a hazard on the road shoulder intruding in the traveling lane. Stationary obstacles used for this test consisted of a single construction barrel and a foam car. Test drivers drove directly toward these obstacles at a specified ACC speed (25 mph, 35 mph, 45 mph) and headway setting (short, medium, long). When the obstacles were approached, the driver was prepared to perform evasive maneuvers in case the system did not recognize the potential for collision and brake accordingly.

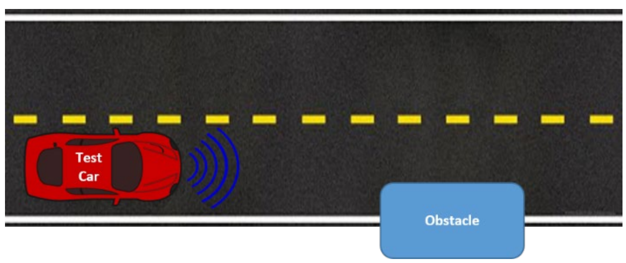


Figure 8. Lane Obstruction scenario diagram.

Temporary Lane Closure

The Temporary Lane Closure test was inspired by several publicly released videos which showed highly automated vehicles colliding into temporary work-zone barriers or construction lane closures. The aim of the test was to examine each vehicle's ACC capabilities when exposed to a temporary lane closure made up of common construction barrel cones. Both full (i.e., whole road) and half lane closures (i.e., single lanes) were considered during testing. Both temporary closures were set up according to VDOT standards, as seen in Appendix D: Additional Test Scenario Images [12].

Lanes used in the test were 10 ft. (3 m) in width. Therefore, for the full lane shift, depending on the vehicle approach speed, the taper length of the lane shift varied from 105 ft. (32 m) to 450 ft. (137.2 m). For the half lane shift, these taper lengths were simply halved. A complete list of the taper lengths used for each specific lane shift iteration can be found in Appendix E: Test Parameters. There were no VDOT specifications or standards for amount of cones to be used in the shift, and therefore cones were placed approximately 10 ft. (3 m) apart to create a smooth lane taper, as seen in Figure 9.

Figure 9. Lane shift scenario diagram (not to scale).

Each test vehicle was exposed to a half lane shift at 25 mph, a full lane shift at 25 mph, a full lane shift at 35 mph, and a full lane shift at 45 mph. Test vehicles were driven centered in the lane directly toward the lane shift at a specified ACC speed and headway setting. The test driver was required to hold the vehicle's central position for as long as possible—until either the vehicle reacted to the construction barrels or the driver was required to perform an evasive maneuver to avoid collision.

Aggregate Automated Feature Testing (Public Road Testing)

To gain a better understanding of how available automated features worked together, test vehicles were driven on the public road with all automated driving systems active. The public road test route, as seen in Appendix D: Additional Test Scenario Images, consisted of a standardized public road circuit that encompassed a variety of road environments and driving scenarios. The highway (e.g., US 460, 460BUS/Christiansburg), urban (e.g., Virginia Tech campus, downtown Blacksburg area), and rural sections of the road (e.g., Blacksburg, Virginia and surrounding Montgomery County areas) allowed researchers to test the automated features in the real-world situations for which they were designed. Features were only engaged according to the manufacturers' specifications while keeping in mind the limitations of each, as determined from the prior controlled testing. Situations on the road were closely monitored by expert drivers and two in-vehicle experimenters to determine if the operational design domains of the features were being exceeded or if the features were not reacting to stimuli as expected. The two in-vehicle researchers also documented vehicle performance and response and/or provided route guidance to drivers. To ensure safety of pedestrians and other road users, automated features were not activated if pedestrians, bicyclists, or other vulnerable road users were present.

Data Analysis

Data obtained from testing were largely observational and qualitative. Narratives describing vehicle responses as well as general observations about tests were recorded by in-vehicle documenters during trials. To convert the qualitative results into data to which statistical analysis could be applied, categorical coding schemes were developed based on vehicle performance during each test. These schemes categorized whether the vehicle had no response, exhibited auditory/visual alerts, attempted mitigation, or demonstrated complete avoidance and/or expected responses to different obstacles or the changing driving landscape. Coding schemes for each test can be found in Appendix F: Coding Schemes.

To aid in eliminating bias during assignment of subjective scores, an inter-rater reliability method was employed. Referencing the specific code scheme assigned to each test, one researcher manually assigned codes to the individual results of each test trial. Separately, another researcher examined the response of the vehicle and independently evaluated each result. The two different sets of scoring were compared, and any discrepancies between code assignments were discussed and resolved. If resolution could not be achieved, project principle investigators were consulted to adjudicate.

Microsoft Excel was used to calculate averages, standard deviations, variances, minimums, maximums, and ranges of the scores for each overall test, specific variable iteration of the test, and individual vehicle. The calculated overall averages demonstrated how vehicles responded to tests on a higher level. Standard deviation was used to better understand the consistency of vehicle performance within each test. From these metrics, researchers were able to obtain a clearer picture of how vehicle responses differed between variable iterations and tests.

Using the calculated metrics, analysis of variance (ANOVA) tests were performed to determine statistically significant differences between test trials. If the ANOVA indicated statistical significance, two tailed t-tests assuming equal variances were performed to determine which variable impacted vehicle performance the most.

Results and Discussion

For the purposes of this report, only overall averages for each test are presented and discussed. In addition, to reduce bias during data analysis and to protect manufacturer identity in publication, vehicle make and model information have been removed and are instead represented numerically in the results. More detailed information and data, including performance scores for specific iterations of parameters and definitions of variables, can be found in the final dataset located on the VTTI Dataverse, referenced in the Data Products section of this report [13]. Results from the statistical tests can be found in Appendix G: Statistical Measures.

ACC Curve

For the ACC curve test, researchers expected the vehicle to continuously track the lead vehicle throughout the entirety of each curve. It was also hypothesized that test vehicles may have moments where the tracking of the lead vehicle would be lost due to the curvature of the turns, which would be indicated visually on the dashboard HMI or by a surge in test vehicle speed. Overall, vehicles scored high, with relatively low standard deviations, as seen in Figure 10. Only one vehicle exhibited behavior that required driver input across all trials.

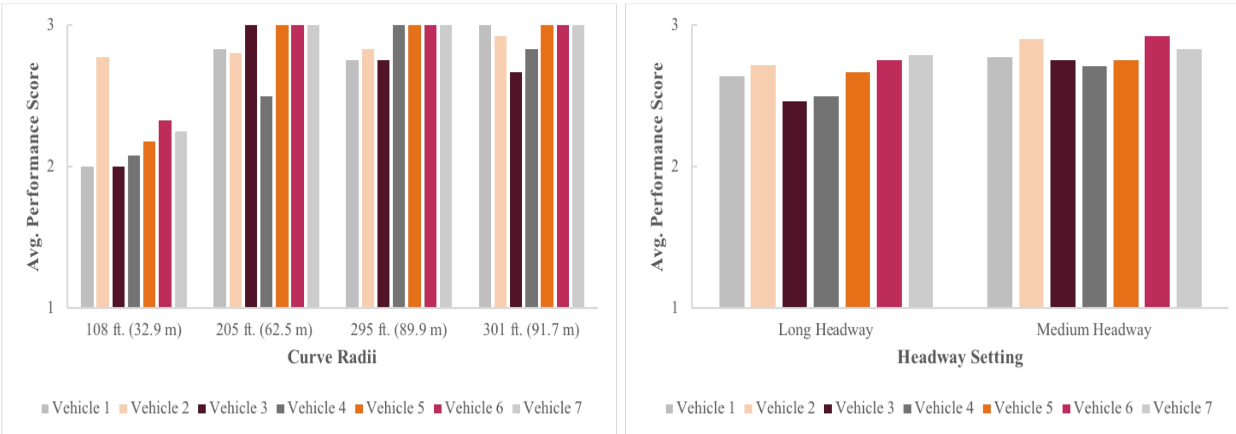


Figure 10. Average performance scores for ACC Curve test based on curve radii (left) and headway (right).

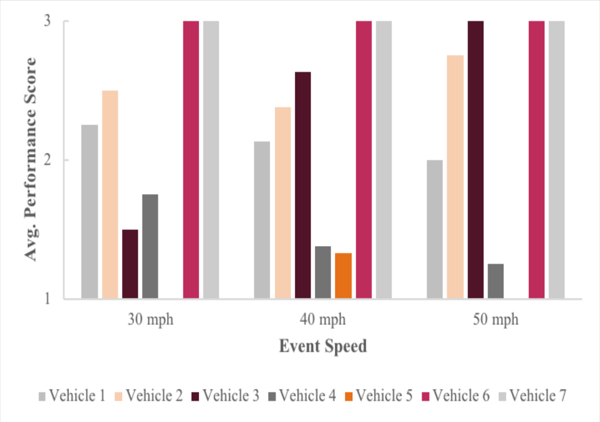
Based on the hypotheses formulated for this test, two ANOVAs were conducted examining average vehicle performance when considering different curve radii and when considering headway setting. Both ANOVAs produced statistically significant results; therefore, t-tests were done to determine which specific curve or headway setting had the largest impact on vehicle performance. From these tests, it was determined that (1) traveling specifically around the 108 ft. (32.9 m) curve onto the Highway section and (2) headway setting had significant impact on vehicle performance.

Vehicles performed slightly worse compared to the overall averages while traveling around the 108 ft. (32.9 m) curve, as this was the only curve where driver intervention was needed. This performance drop could be due to the rate of curvature. Since the angle of the curve is more severe than the other curves, it is possible that the sensor field of views could not continually capture and track the lead vehicle [14, 15]. The other curves on the Highway have much larger radii, which gave the vehicle more time to adjust and continue tracking the lead vehicle. As mentioned previously, headway setting also had an effect on vehicle performance.

All vehicles had higher average performance scores while the medium headway setting was engaged. When ACC was set to a medium headway, the test vehicle more closely followed the lead vehicle. Therefore, around curves, it is possible that the lead vehicle remained in the sensor and camera field of view longer, potentially not losing track of the lead vehicle at any point in the curves. Continuous lead vehicle tracking will increase the test vehicle performance, as the test vehicle will be able to adjust speed and following distance accordingly. In addition, if the test vehicle continuously tracks the lead vehicle, there is lowered risk of experiencing unexpected speed surges, which could require driver takeover without much warning [14, 15].

ACC Cut-in

Although ACC-equipped vehicles performed well during previous studies which examined vehicles fully cutting into the traveling lane, when the cut-in vehicle only partially cut into the traveling lane, the test vehicles seemed to have a difficult time recognizing the lead vehicle [16]. Vehicles exhibited satisfactory performance, as seen in Figure 13, but with high standard deviations. Three out of the seven test vehicles required driver input at some point during the tests. In certain instances, the test vehicle did not recognize the cut-in vehicle as the new lead vehicle and did not appropriately adjust speed nor following distance, requiring the test vehicle drivers to take over. As exhibited on the vehicle's visual HMI systems, there was a lot of "bouncing back and forth" between the detection of the two lead vehicles. This "bouncing" could be an indication of test vehicle confusion in identifying the "true" lead vehicle, which caused multiple surges in speed during tests.



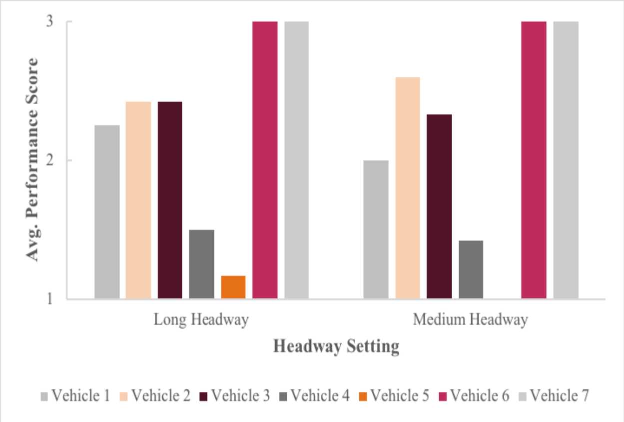


Figure 11. Average performance scores for ACC Cut-In test considering event speeds (left) and vehicle headway setting (right).

Performance differences due to both event speeds and headway settings were examined through ANOVA tests. Although neither ANOVA showed statistically significant effects on test vehicle performance due to either of these factors, differences in test vehicle performance across the different speed variations were observed by in-vehicle researchers. As noted previously, three out of seven test vehicle drives required intervention and manually speed adjustments or following distance to avoid collision with the cut-in vehicle. One vehicle even needed driver intervention for every single test iteration.

One possible explanation for the difficulty of lead vehicle detection could be the high variability of the cut-in vehicle's position. For some vehicles, if the cut-in vehicle was not perfectly aligned on the center lane line, the test vehicle would not identify it as the new lead vehicle and subsequently not appropriately adjust speed or following distance. However, if the vehicle was perfectly centered on the lane line or slightly encroaching in the traveling lane, the test vehicle identified it as the new lead vehicle and adjusted accordingly. This variability in adjustment points to potential issues with limited field of view in both the cameras and sensors on these vehicles.

ACC Cut-out

As seen in Figure 15, for tests with revealed vehicles of 15 mph and 10 mph, a majority of the test vehicles performed almost perfectly. In these scenarios, vehicles adjusted to the new revealed vehicle speed accordingly, without any driver intervention. In addition, these tests had low standard deviation, consistently meeting performance expectations. However, vehicles did not exhibit the same type of expected responses to stationary revealed vehicles.

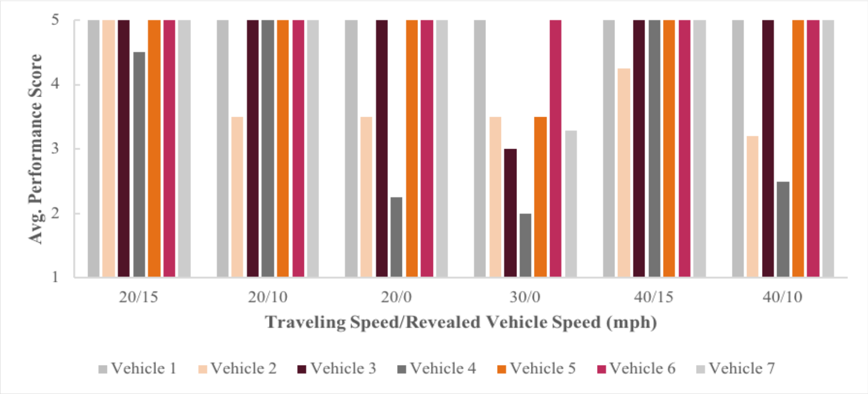


Figure 12. Average performance scores for ACC Cut-Out test based on differences between traveling vehicle speed and revealed vehicle speed.

ANOVA tests were performed to examine the different event speed combinations and ACC headway settings. The ANOVA test comparing the headway settings did not produce any significant results, and therefore it can be suggested that headway setting did not have an effect on vehicle performance during these tests. However, the ANOVA test examining the different event speeds produced statistically significant results. T-tests were therefore done to identify which combination of traveling speed and revealed vehicle speed had the largest impact on vehicle performance. The only t-tests that produced significant results were those which had revealed speeds of 0 mph, specifically those where the traveling speed was 30 mph. Because significance was shown, it can be concluded that the revealed vehicle speed was the main factor that influenced the test vehicle fleet performance during the ACC Cut-out test.

Qualitatively, in trials with revealed speeds of 15 mph and 10 mph, vehicles consistently performed in alignment with researchers' initial expectations. In some instances, when the revealed vehicle was traveling at 10 mph, two out of the seven vehicles did not apply any sort of braking, or applied only mild braking, to avoid the vehicle in its path, instead exhibiting warning lights and sounds to alert the driver of the collision risk.

The tests in which the revealed vehicle was stationary had low performance scores relative to the other trials. However, all vehicles still exhibited some sort of response to the obstacle in the driving path. These tests may have had such poor performance since ACC is not specifically designed to handle high-speed difference situations and has been proven in the past to not adequately detect stationary objects [17, 18, 19, 20, 21, 22, 23]. This type of scenario may be better suited for testing AEB systems, for which some vehicles began to activate before the driver intervened and braked during testing.

Stop and Go

In this test, vehicles had consistent high-performance scores, as seen in Figure 16. No driver input was needed during these tests due to proper test vehicle response to the lead vehicles.

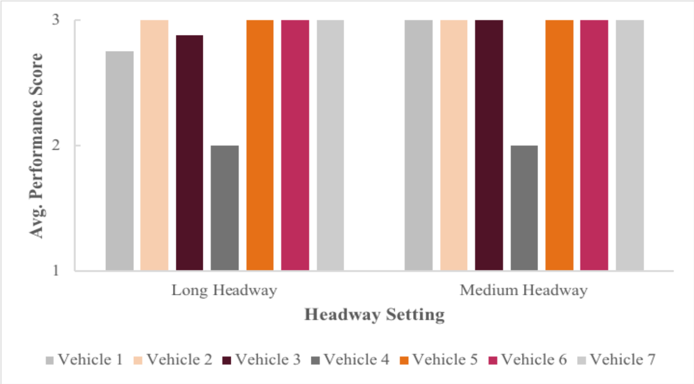


Figure 13. Average performance scores for Stop and Go test.

In this test, headway settings were the only variable manipulated, and therefore headway was the only consideration during statistical analysis. The test concluded that there was no statistical difference in vehicle performance due to headway setting. In addition, the high level of vehicles' performance in this test is not surprising, as most automated vehicle features are designed specifically for dynamic speeds while following a singular lead vehicle. However, for the last speed change during the test (i.e., 0 mph to 30 mph) some test vehicles' ACC needed to be re-engaged manually in order to continue following the lead vehicles.

LKA Inattentiveness

Because the LKA Inattentiveness test was highly specific to each vehicle type and researchers were not examining the same types of responses as in the other tests, results were not given specific performance scores as in the tests above. Instead, the time taken to reach a warning stage (i.e., auditory warning, visual warning, or disengagement of the automated system) was evaluated. The timing for these warnings was considered as the main variable for analysis. Time, in seconds, from hands off the wheel to in-vehicle warnings and maneuvers can be seen in Table 2.

Table 2. Time to In-Vehicle Warnings (Seconds)

	Primary	Secondary	Disengagement
Overall	16.44	20.75	22.64
Vehicle 1	47.50	61.50	107
Vehicle 2	11.01	15.83	20.60
Vehicle 3	15.09	25.43	18.17
Vehicle 4	None	None	15.04
Vehicle 5	23.43	None	25.70
Vehicle 6	42.50	56.75	None
Vehicle 7	None	None	None

Primary warnings were the first warnings that the vehicle emitted after the test drivers' hands were off of the wheel for a certain amount of time. Typically, these warnings were visual and appeared on the dashboard HMI. Secondary warnings, emitted if no action was taken from the primary warning, were typically auditory warnings. Disengagement was classified as when the automated system (e.g., ACC, LKA) disengaged, forcing the driver to regain control. It is important to note that vehicles exhibited quite different time-to-warnings in this test. Some vehicles displayed warnings at similar time intervals across tests, while others took almost double the time to show warnings or even failed to provide any sort of warning before automated feature disengagement.

AEB Obstacle

In this test, AEB was the only feature evaluated; therefore during this experiment, none of the ACC features were active. Test vehicles ended up scoring low (i.e., no stopping or warnings) and demonstrated inconsistent responses to the roadway obstacles, as seen in Figure 17. In addition, this test had some of the highest standard deviations of any of the tests conducted with the test vehicle fleet.

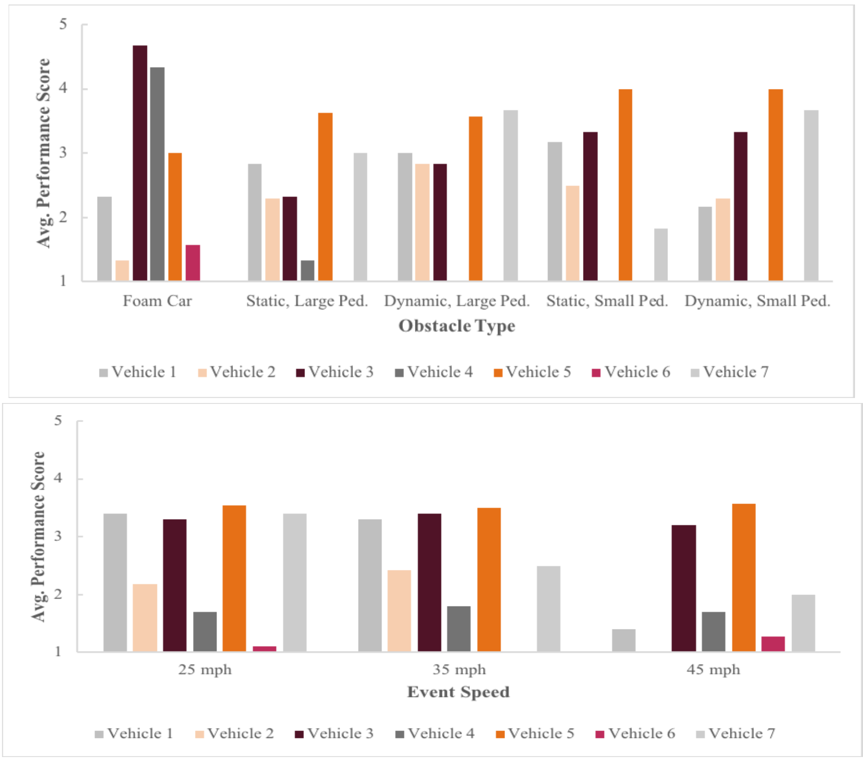


Figure 14. Average performance scores for AEB Obstacle test based on obstacle type (top) and event speed (bottom).

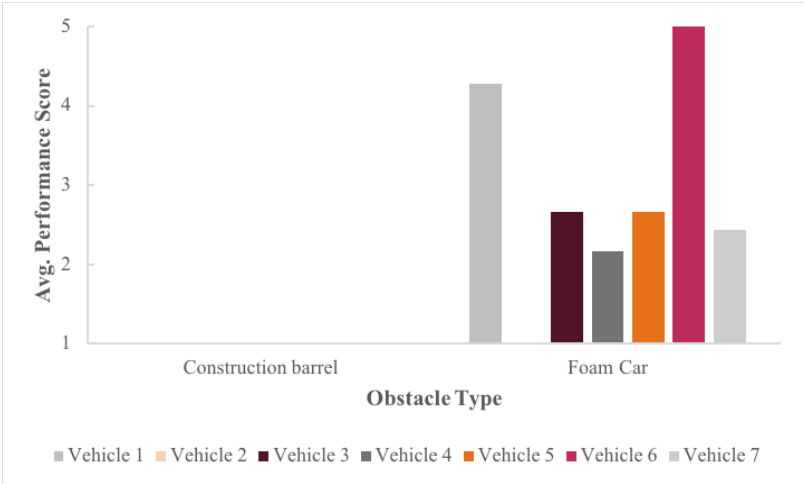
ANOVAs were carried out to examine vehicle performance based on obstacle type and traveling speed, which indicated no statistical significance. Although no statistical significance was found, based on summary data, it appears that larger targets (e.g., foam car and large pedestrian target) and approaching obstacles at lower speeds (e.g., 25 mph) resulted in the highest performance scores.

Surprisingly, the small pedestrian target produced high levels of vehicle response relative to the other obstacle results. One vehicle detected the dynamic small pedestrian target better than any other obstacle. This result is unexpected since most vehicle manuals specifically state that smaller pedestrians, such as children, will not be recognized by the automated system's sensors and cameras [17, 18, 19, 20, 21, 22, 23].

All vehicles tested were equipped with AEB systems, which were active at the time of testing. Therefore, it is surprising that a majority of the test vehicles did not consistently react to the targets in the expected manner. One possible explanation for the lack of vehicle response could be the validity of the targets used as obstacles. These targets are mostly utilized for VTTI in-house experimentation and have not been externally validated, which could make them inadequate for this type of experimentation. For example, the soft targets may not reflect a radar signature or visual features that would produce the intended detection by radar and camera sensors. However, because some vehicles reacted to the obstacles, just inconsistently, this behavior could also suggest an opportunity for improvement in L1/L2 and AEB capabilities.

Lane Obstruction

Vehicles scored relatively low in this test, as seen in Figure 19, with none of the test vehicles perceived nor reacting to the construction barrel in their path. Even the foam car, which vehicles responded effectively to in the AEB tests, only evoked responses from vehicles in about half of all trials. In addition, these tests produced results with higher standard deviations compared to other tests. Based on the consistent lack of vehicle response to the construction barrels, these results were excluded from statistical data analysis, specifically when examining speed and headway interactions.



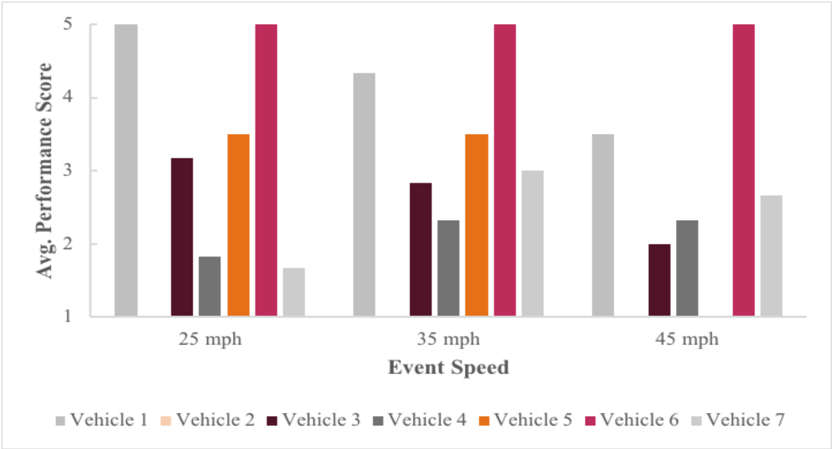


Figure 15. Average performance scores from lane obstruction test (maximum score of 5) by obstacle type (top) and event speed (bottom).

ANOVAs were performed to examine the effects that the two different obstacles, the traveling speeds, and the headway settings had on vehicle performance. Neither traveling speeds nor headway settings had significant impacts on vehicle performance. The ANOVA indicated there was a statistical difference between vehicle performances when exposed to the two different objects, since vehicles consistently exhibited no response to the construction barrels.

Considering the results obtained from this test, vehicles exhibited neither consistent nor expected responses to any of the presented obstacles. This performance deficiency could be due to the vehicles' ACC capabilities, as ACC is not specifically designed to handle situations with stationary objects, as stated in vehicle manuals. On the other hand, specifically in the instances of trials that used the construction barrel, the results indicate that current automated vehicle sensors, cameras, and/or machine vision software may not be equipped to detect smaller, potentially safety-critical obstacles, which could pose threats to vulnerable road users (e.g., people wearing reflective vests/using reflective construction barrels such as in construction zones) or in emergency situations.

Lane Shift

Similar to the performance seen in the Lane Obstruction test construction barrel trials, a majority of the vehicles performed relatively poorly in this test, as shown in Figure 20. Four out of the seven vehicles had no response to the construction barrel lane shift at any speed or shift angle iteration. Only one vehicle consistently reacted to the lane shift by exhibiting warning lights and/or sounds. Two other vehicles inconsistently reacted to the lane shifts with warning lights and/or sounds.

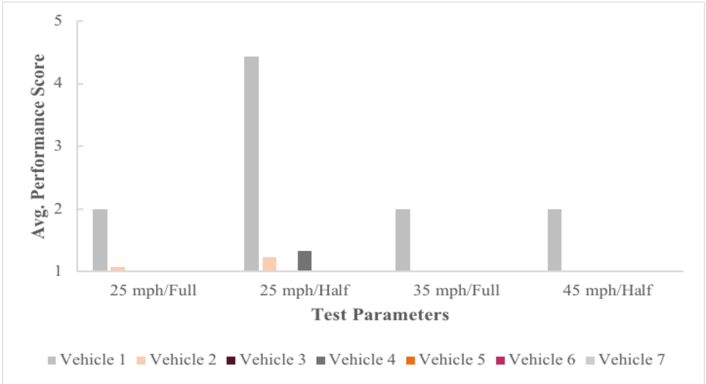


Figure 16. Average performance scores from Lane Shift Test (maximum score of 5; traveling speed/shift severity).

ANOVA tests were performed to examine the effect of lane shift type and headway setting on vehicle performance. Based on the test results, there was no statistical difference found between vehicle performance and these metrics.

However, the 25-mph half lane shift evoked the most vehicle responses and response intensity. In this configuration, three vehicles reacted to the shift, and one even came to a full stop right before the shift began. From a passenger's perspective, at a distance, this lane shift looked almost like a solid wall due to the arrangement of the cones. Since this shift looked more like a solid wall, vehicle cameras and/or sensors may have recognized it as such and engaged AEB. Given the number of sensors and cameras that came equipped on these test vehicles, researchers hypothesized that the vehicles should be capable of detecting the temporary barriers. However, this test demonstrated there are opportunities for improvement to be made to current automated vehicle detecting capabilities, especially for stationary obstacles.

Aggregate Automated Feature Testing

The public test route was broken up into four road classes: highway, urban, and rural, as described in the Methods section. In addition, based on the frequency with which they occurred, key events were identified. Performance scores were given based on the vehicle's ability to handle these daily driving scenarios, as shown in the table located in Appendix xxx. Not all events were experienced during every drive; gray boxes in the table below indicate events that did not take place in particular driving scenarios.

Not all vehicles experienced the same type of events or the same number of events during a trip. However, the results seen from this test confirm the results that were also derived from the Smart Road tests. Vehicles seemed to perform better when following lead vehicles or adapting to stop and go traffic, as these situations are more within the operational design domain of ACC compared to reacting to stationary obstacles in the roadway.

Conclusions and Recommendations

From the tests conducted in this study, it was determined that:

1. Curve severity had an effect on test vehicle following performance while ACC was engaged.
2. Headway setting had an effect on test vehicle following performance while ACC was engaged.
3. Test vehicle ACC performance was affected when a revealed vehicle was stationary.
4. Test vehicles were not able to sense small obstacles, such as construction barrels.
5. Test vehicles exhibited inconsistent responses to identical configurations of stimuli.

Overall, a majority of the test vehicles exhibited a response (e.g., warning lights/sounds) to obstacles in their path; however, most of the time the vehicle did not brake or reduce speed to avoid or mitigate collision. The vehicles were also inconsistent in their performance, which demonstrated the large variations in test fleet vehicle capabilities even though, theoretically, they should perform at similar levels given their L2 capabilities. In addition, not only did each vehicle's performance capabilities vary, but the physical manner in which the automated features were engaged and by which they communicated their status to the driver differed considerably between each vehicle, which could create confusion for a user.

Through this project, an initial framework of standardized tests was developed to evaluate automated vehicle capabilities in real-world scenarios. In addition to the basic testing framework, researchers were able to use these tests to create baseline results, which appear valid for currently advertised automated vehicles, and gain a better understanding of the discrepancies and variations of features currently in these vehicles.

Testing limitations may have contributed to the inconsistency in vehicle performance. For example, as mentioned in the Results section, some tests were not customized where specific vehicles had specific feature limitations. For some tests, the vehicles approached a stationary target with ACC engaged; however, many of the vehicle manuals stated that ACC will not work appropriately "if the lead object is stationary." In addition, tests such as AEB Obstacle and Lane Obstruction were performed with targets that were not externally validated and therefore may not have been accurate representations of actual objects.

However, this project did achieve the goal of developing a framework of standardized testing that can continue to be refined and applied to new vehicles as the performance of vehicle automation improves. Some of the tests conducted were designed considering more advanced vehicles than the ones that used for this study, which should allow the measures to be sensitive to future capabilities. The study has shown that low-cost standardized testing can be applied to ADAS with sufficient trends to indicate meaningful characterization of relative performance. Anecdotally, the observational measures and statistical outcomes obtained match the researchers' expectations.

Indeed, it appears that this testing method has value and could form a basis for more robust standardized testing, possibly as part of a self- or public- certification program.

Additional Products

The Education and Workforce Development (EWD) and Technology Transfer (T2) products created as part of this project are described below and are listed on the Safe-D website here. The final project dataset is located on the Safe-D Dataverse.

Education and Workforce Development Products

This project allowed students to take on leadership roles and learn about conducting high-fidelity vehicle research throughout all stages of the process. Students were heavily involved in the initial grant proposal and performed extensive literature reviews in the subject area. In addition, they took on the primary responsibilities of designing the Smart Road experiment and making sure it adhered to VTTI safety policies. Finally, students organized and conducted all Smart Road tests to collect data.

Outside of development and execution of the study, students also gained experience writing academic journal papers and going through the peer-review process. An abstract on the study was submitted to the National Highway Traffic Safety Administration Enhanced Safety of Vehicles conference and went through the peer-review process for the *Journal of Traffic Injury Prevention*. Ultimately, an academic journal paper about the study was published in the SAE International *Journal of Connected and Automated Vehicles* [24]. In addition, students gained valuable public speaking skills by presenting the work at numerous conferences and events.

Technology Transfer Products

As mentioned previously, this project produced an academic journal paper that was published in the SAE International *Journal of Connected and Automated Vehicles*. In addition, a podium talk was given about the study at the 2019 FAST-zero Conference in Blacksburg, Virginia. These T2 products are listed on the project page of the Safe-D website.

Data Products

The data uploaded to the Dataverse includes all raw subjective data recorded during testing by in-vehicle experimenters, vehicle performance descriptions, and statistical analyses performed. Data from this project is available from the VTTI Safe-D Dataverse listed under project #VTTI-00-020. The dataset can be accessed at:
https://dataverse.vtti.vt.edu/dataset.xhtml?persistentId=doi:10.15787/VTT1/D946JJ

References

1. Pfeffer, R., & Leichsenring, T. (2016). Continuous Development of Highly Automated DCriving Functions with Vehicle-in-the-Loop Using the Example of Euro NCAP Scenarios. *Simulation and Testing for Vehicle Technology*, 33-42.
2. Terry, T., & Tanner, S. (2017). *Problematic Roadway Enviornments for Automated Vehicles- Database Development.* National Surface Transportation Safety Center for Excellence.
3. Sippl, C., Bock, F., Whittmann, D., Altinger, H., & German, R. (2016). From Simulation Data to Test Cases for Fully Automated Driving and ADAS. *ICTSS*, 191-206.
4. SAE International. (2018). *J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.*
5. NHTSA. (2017). Automated Vehicles for Safety. https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety
6. Sternlund, S., Strandroth, J., Rizzi, M., Lie, A., & Tingvall, C. (2017). *The effectiveness of lane departure warning systems- A reduction in real-world passenger car injury crashes.* Traffic Injury Prevention.
7. Agaram, V., Barickman, F., Fahrenkrog, F., Griffor, E., Muharemovic, I., Peng, H., . . . Shogren, W. (2016). Validation and Verification of Automated Road Vehicles. *Road Vehicle Automation 3*, 201-210.
8. Knauss, A., Schroder, J., Berger, C., & Eriksson, H. (2017). Paving the roadway for safety of automated vehicles: An empirical study on testing challenges. *IEEE Intelligent Vehicles Symposium* , 1873-1880.
9. Zhao, D., & Peng, H. (2017). Solving the Challenge of Accelerating Automated Vehicle Testing. *Mcity White Paper*.
10. IIHS. (2018). *Reality check: Research, deadly crashes show need for caution on the road to full autonomy.*
11. Euro NCAP. (2018). *Automated Driving Tests.*
12. Virginia Department of Transportation. (2012). Work Zone Safety: Guidelines for Temporary Traffic Control. In *2011 Virginia Work Area Protection Manual.*
13. Basantis, A., Doerzaph, Z., Harwood, L., & Neurauter, L. (2019). *Standardized Performance Evaluation of Vehicles with Automated Capabilities (VTTI-00-020).* VTTI. doi:https://doi.org/10.15787/VTT1/D946JJ

14. International Organization for Standardization. (2010). *ISO 15622: Intelligent transport systems- Adaptive cruise control systems- Performance requirements and test procedures.*

15. International Standard Organization. (2009). *ISO 22179: Intelligent transport systems- Full speed range adaptive cruise control (FSRA) systems- Performance requirements and test procedures.*

16. Milanes, V., Shladover, S., Spring, J., Nowakowski, C., Hiroshi, K., & Nakamura, M. (2014). Cooperative Adaptive Cruise Control in Real Traffic Situations. *IEEE Transactions on Intelligent Transportation Systems, 15*(1), 296-305.

17. Audi AG. (2017). *Q7 Owner's Manual.*

18. Daimler AG. (2016). *Mercedes-Benz E-Class Sedan and Wagon Owner's Manual.*

19. General Motors LLC. (2018). *Cadillac CT6 Owner's Manual.*

20. Infiniti. (2015). *Q50 Owner's Manual.*

21. Tesla. (2015). *Model S Owner's Manual.*

22. Tesla. (2018). *Model X Owner's Manual.*

23. Volvo Cars LLC. (2016). *XC 90 Owner's Manual.*

24. Basantis, A., Doerzaph, Z., Harwood, L., and Neurauter, L., "Developing a Standardized Performance Evaluation of Vehicles with Automated Driving Features," SAE Intl. J CAV 2(3):2019, https://doi.org/10.4271/12-02-03-0011.

Appendix

Appendix A: Focus Group Questions

1. What low-level autonomous feature (ACC, AEB, Lane keeping) do you think provides the least room for error? (I.e. system failure will cause the most danger/harm)
2. Do AEB systems on today's vehicles have more difficulty with stationary or moving objects?
3. Are there common cases of AEB systems incorrectly identifying a dangerous situation and applying the brakes? (Such as a collision with a flying plastic bag, bird, or traffic cone)
4. Are there any specific scenarios you would like to see on a VTTI automated features testing standard?
5. Are there concerns about feature failures associated with light wash-out on camera systems? (driving directly into sun, oncoming car with high beams)
6. Do you think any meaningful knowledge can be gained from testing low-level autonomous systems (ACC, AEB, Lane keeping) outside their operational limits when it is known they will fail?

Appendix B: Fleet Vehicles and Automated Driver Assistance System (ADAS) Packages

Vehicle	ACC	LKA	AEB
2017 Audi Q7	Audi Adaptive Cruise Control	Audi Active Lane Assist	Audi Pre Sense
2015 Infiniti Q50	Intelligent Cruise Control	Active Lane Control	Forward Emergency Braking
2016 Mercedes-Benz E350	Distronic Plus	Not specified	Brake Assist System Plus, Collision Prevention Assist Plus, Pre-Safe Brake
2015 Tesla Model S	Traffic-Aware Cruise Control	Lane Departure Warning	Forward Collision Warning
2016 Volvo XC90	Adaptive Cruise Control, Pilot Assist	Pilot Assist, Lane Departure Warning, Lane Keeping Aid	Pilot Assist, City Safety
2018 Cadillac CT6	Adaptive Cruise Control	Lane Keep Assist with Lane Departure Warning	Low Speed Forward Automatic Braking, Front Pedestrian Braking
2018 Tesla Model X	Traffic-Aware Cruise Control	Lane Departure Warning	Forward Collision Warning

Appendix C: Expected Vehicle Performance

Based on the different conditions tested, hypotheses regarding vehicle performance were formulated. These hypotheses also applied to statistical testing done on experimental data. Both null hypotheses (H_{0x}) and alternate hypotheses (H_x) were formulated for each experimental test.

ACC Curve

H_{o1}: Curve type does not affect test vehicle performance.
H_1: Curve type affects test vehicle performance.
H_{o2}: Headway setting affects test vehicle performance while maneuvering curves with ACC engaged.
H_2: Headway setting does not affect test vehicle performance while maneuvering curves with ACC engaged.

ACC Cut-In

H_{o3}: Speed traveled does not affect test vehicle performance.
H_3: Speed traveled affects test vehicle performance.
H_{o4}: Headway setting does not affect test vehicle performance during partial cut-in maneuvers.
H_4: Headway setting affects test vehicle performance during partial cut-in maneuvers.

ACC Cut-Out

H_{o6}: Revealed vehicle speeds do not affect test vehicle performance.
H_6: Revealed vehicle speeds affect test vehicle performance.
H_{o7}: Headway setting does not affect test vehicle performance during cut-out maneuvers.
H_7: Headway setting affects test vehicle performance during cut-out maneuvers.

Stop and Go

H_{o8}: Headway setting does not affect test vehicle performance.
H_8: Headway setting affects test vehicle performance.

AEB Obstacle

H_{o9}: Obstacle type does not affect test vehicle AEB performance.
H_9: Obstacle type affects test vehicle AEB performance.
H_{o10}: Speed traveled toward obstacles does not affect test vehicle AEB performance.
H_{10}: Speed traveled toward obstacles affects test vehicle AEB performance.

Lane Obstruction

H_{o11}: Obstacle type does not affect test vehicle performance while ACC is engaged.
H_{11}: Obstacle type affects test vehicle performance while ACC is engaged.
H_{o12}: Headway setting does not affect test vehicle performance.
H_{12}: Headway setting affects test vehicle performance.
H_{o13}: Speed traveled toward obstacle does not affect test vehicle performance.
H_{13}: Speed traveled toward obstacle affects test vehicle performance.

Lane Shift

H_{o14}: Headway setting does not affect test vehicle performance.
H_{14}: Headway setting affects test vehicle performance.
H_{o15}: Shift severity and traveling speed does not affect test vehicle performance.
H_{15}: Shift severity and traveling speed affects test vehicle performance.

Appendix D: Additional Test Scenario Images

ACC Curve

Lane Obstruction

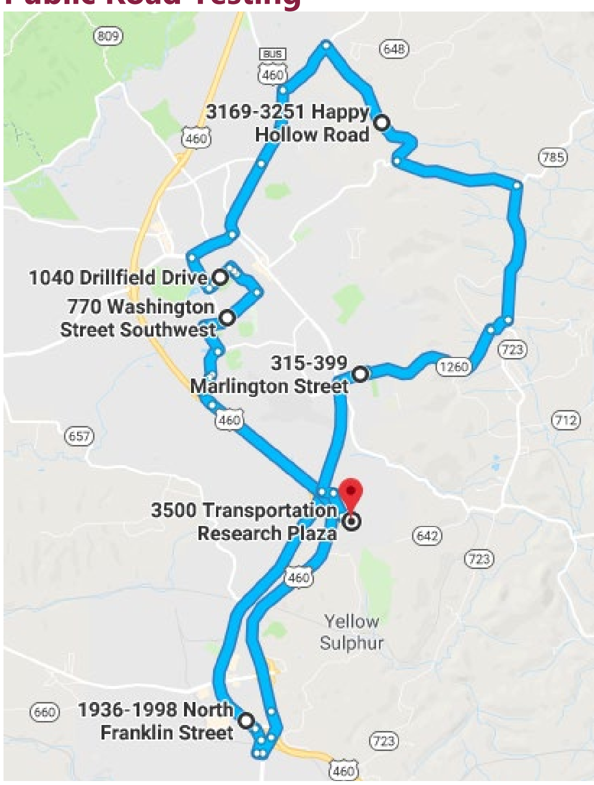
Lane Shift

Taper Length (L) Chart:

Taper Length (L)					
Speed (S) in MPH	Width of Offset (W) in Ft.				Remarks
	9	10	11	12	
25 or below	95	105	115	125	$L=S^2W/60$
30	135	150	165	180	"
35	185	205	225	245	"
40	240	270	295	320	"
45	405	450	495	540	L=SW
50	450	500	550	600	"
55	495	550	605	660	"
60	540	600	660	720	"
65	585	650	715	780	"
70	630	700	770	840	"

Minimum Merging Taper Length (L) for Limited Access Highways Shall Be 1000 Feet. Minimum Shifting Taper Shall Be ½L with L Desirable.

Public Road Testing



Appendix E: Test Parameters

ACC Curve

Parameter	Iterations
Lead Vehicle Speed	15 mph, 20 mph, 25 mph
Test Vehicle ACC Set Speed	20 mph, 25 mph, 30 mph
ACC Headway Setting	Medium, long
Curve Radius	108 ft. (32.9 m), 205 ft. (62.5 m), 295 ft. (89.9 m), 301 ft. (91.7 m)

ACC Cut-In

Parameter	Iterations
Lead Vehicle 1 Speed	30 mph, 40 mph, 50 mph
Lead Vehicle 2 Speed (Cut-In Vehicle)	30 mph, 40 mph, 50 mph
Test Vehicle ACC Set Speed	35 mph, 45 mph, 55 mph
ACC Headway Setting	Medium, long

ACC Cut-Out

Parameter	Iterations
Test Vehicle ACC Set Speed	25 mph, 35 mph, 45 mph
Lead Vehicle Speed	20 mph, 30 mph, 40 mph
Revealed Vehicle Speed	0 mph, 10 mph, 15 mph
ACC Headway Setting	Medium, long

Stop and Go

Parameter	Iterations
ACC Headway Setting	Medium, long

LKA Inattentiveness

Parameter	Iterations
Test Vehicle Speed	45 mph, 55 mph
Steering Correction Setting	Vehicle specific (may or may not be present on all vehicles)

AEB Obstacle

Parameter	Iteration
Test Vehicle Speed	25 mph, 35 mph, 45 mph
Obstacle Type	Large static pedestrian, large moving pedestrian, small static pedestrian, small moving pedestrian, foam car

Lane Obstruction

Parameter	Iterations
Test Vehicle Speed	25 mph, 35 mph, 45 mph
ACC Headway Setting	Short, medium, long
Obstacle Type	Foam car, construction barrel

Lane Shift

Parameter	Iterations
Test Vehicle Speed	25 mph, 35 mph, 45 mph
Lane Reroute Severity	Half lane width, full lane width
ACC Headway Setting	Short, medium, long

Speed (mph)	Shift Width	Length of Taper (ft.)	Length of Taper (m)
25	Full (10 ft.)	105	32
25	Half (5 ft.)	52.5	16
35	Full (10 ft.)	205	62.5
45	Half (5 ft.)	225	68.6

Appendix F: Coding Schemes

ACC Curve

Results Key	
Value	**Description**
3	No driver intervention needed No false positives or failure of system
2	Intermittent system failure No driver intervention needed
1	Driver required to input lateral or longitudinal control to avoid collision

ACC Cut-In

Results Key	
Value	**Description**
3	No driver intervention needed No false positives or failure of system
2	Intermittent system failure No driver intervention needed
1	Driver required to input lateral or longitudinal control to avoid collision OR no response

ACC Cut-Out

Results Key	
Value	**Description**
5	No driver intervention needed
4	Driver required to input lateral control or braking to avoid collision Car performed hard self-braking Warning lights and sounds
3	Driver required to input lateral control Car performed mild self-braking Warning lights or sounds
2	Warning lights or sounds
1	No car response to event

Stop and Go

Results Key	
Value	**Description**
3	No driver intervention needed No false positives or failure of system
2	Intermittent system failure No driver intervention needed
1	Driver required to input lateral or longitudinal control to avoid collision OR no response

AEB Obstacle

Results Key	
Value	**Description**
5	No driver intervention needed
4	Driver required to input lateral control or braking to avoid collision Car performed hard self-braking Warning lights and sounds
3	Driver required to input lateral control Car performed mild self-braking Warning lights or sounds
2	Warning lights or sounds
1	No car response to event

Lane Obstruction

Results Key	
Value	**Description**
5	No driver intervention needed
4	Driver required to input lateral control or braking to avoid collision Car performed hard self-braking Warning lights and sounds
3	Driver required to input lateral control Car performed mild self-braking Warning lights or sounds
2	Warning lights or sounds
1	No car response to event

Lane Shift

Results Key	
Value	**Description**
5	No driver intervention needed
4	Driver required to input lateral control or braking to avoid collision Car performed hard self-braking Warning lights and sounds
3	Driver required to input lateral control Car performed mild self-braking Warning lights or sounds
2	Warning lights or sounds
1	No car response to event

Aggregate Feature Testing

Results Key	
Value	**Description**
3	No driver intervention needed No false positives or failure of system

2	Intermittent system failure No driver intervention needed
1	Driver required to input lateral or longitudinal control to avoid collision

Appendix G: Statistical Measures

Highlighted values indicate significance.

ACC Curve

Curve Type

Groups	Average	Variance
108 ft. (32.9 m)	2.23	0.07
301 ft. (91.7 m)	2.92	0.02
295 ft. (89.9 m)	2.90	0.01
205 ft. (62.5 m)	2.88	0.03

ANOVA	
F	F-crit
22.91	3.01

T-Tests		
Curve Severity 1	Curve Severity 2	P-value
108 ft. (32.9 m)	301 ft.	5.15E-05
108 ft. (32.9 m)	295 ft.	5.80E-05
108 ft. (32.9 m)	205 ft.	0.0002
301 ft. (91.7 m)	295 ft.	0.850
301 ft. (91.7 m)	205 ft.	0.670
295 ft. (89.9 m)	205 ft.	0.778

Headway

Groups	Average	Variance
Long headway	2.65	0.02
Medium headway	2.81	0.01

ANOVA	
F	F-crit
7.83	4.75

ACC Cut-In

Event Speed

Groups	Average	Variance
30 mph	2.14	0.58
40 mph	2.26	0.48
50 mph	2.29	0.76

ANOVA	
F	F-crit
0.07	3.55

Headway

Groups	Average	Variance
Long headway	2.23	0.07
Medium headway	2.92	0.02

ANOVA	
F	F-crit
0.02	4.75

ACC Cut-Out

Traveling (TS) and Revealed (RS) Vehicle Speeds

Groups	Average	Variance
TS: 20 mph, RS: 15 mph	4.93	0.04
TS: 20 mph, RS: 10 mph	4.79	0.32
TS: 20 mph, RS: 0 mph	4.39	1.21
TS: 30 mph, RS: 0 mph	3.61	1.16
TS: 40 mph, RS: 15 mph	4.89	0.08
TS: 40 mph, RS: 10 mph	4.39	1.14

ANOVA	
F	F-crit
2.62	2.48

T-Tests		
TS/RS 1 (mph)	TS/RS 2 (mph)	P-value
20/15	20/10	0.27
20/15	20/0	0.23
20/15	30/0	0.008
20/15	40/15	0.79
20/15	40/10	0.21
20/10	20/0	0.42
20/10	30/0	0.03
20/10	40/15	0.66
20/0	30/0	0.20

Headway

Groups	Average	Variance
Long headway	4.73	0.27
Medium headway	4.23	0.73

ANOVA	
F	F-crit
1.80	4.75

Stop and Go

Headway

Groups	Average	Variance
Long headway	2.80	0.13
Medium headway	2.86	0.14

ANOVA	
F	F-crit
0.07	4.75

AEB Obstacle

Obstacle Type

Groups	Average	Variance

Foam Car	2.60	2.12
Stationary Large Pedestrian	2.34	0.86
Dynamic Large Pedestrian	2.56	1.24
Stationary Small Pedestrian	2.40	1.38
Dynamic Small Pedestrian	2.49	1.49

ANOVA	
F	**F-crit**
0.06	2.69

Traveling Speed

Groups	Average	Variance
25 mph	2.66	0.98
35 mph	2.56	0.86
45 mph	2.13	0.80

ANOVA	
F	**F-crit**
0.62	3.55

Lane Obstruction

Obstacle Type

Groups	Average	Variance
Construction barrel	1	0
Foam Car	2.89	1.80

ANOVA	
F	**F-crit**
13.93	4.75

Headway

Groups	Average	Variance
Long headway	3.12	2.17
Medium headway	2.99	2.14
Short headway	2.72	1.82

ANOVA	
F	**F-crit**
0.14	3.55

Traveling Speeds

Groups	Average	Variance
25 mph	3.02	2.57
35 mph	3.14	1.73
45 mph	2.50	2.01

ANOVA	
F	**F-crit**
0.39	3.55

Lane Shift

Headway

Groups	Average	Variance

Long headway	1.25	0.39
Medium headway	1.27	0.37
Short headway	1.28	0.42

ANOVA	
F	**F-crit**
0.005	3.55

Traveling Speed/Shift Type

Groups	Average	Variance
25 mph, Full Shift	1.15	0.14
25 mph, Half Shift	1.57	1.61
35 mph, Full Shift	1.17	0.17
45 mph, Half Shift	1.14	0.14

ANOVA	
F	**F-crit**
0.57	3.03

Appendix F: Aggregate Feature Testing Results

Aggregate Feature Testing Events, Scores, and Number of Observations (n)

Vehicle	Following Lead Vehicle	Lead Vehicle Slowing	Maintain Speed-Downhill	Maintain Speed-Curves	Maintain Speed-Traffic	Stop and Go	Pedestrian Detection	Stationary Object	Vehicle Cut-In
Vehicle 1	3 (n=1)		3 (n=1)	1.75 (n=3)		3 (n=1)			
Vehicle 2	2.5 (n=2)		3 (n=1)	1 (n=2)		3 (n=1)	2 (n=1)	1.29 (n=7)	
Vehicle 3	3 (n=1)	2 (n=1)	3 (n=1)	3 (n=2)		3 (n=2)		3 (n=5)	
Vehicle 4	3 (n=2)	2 (n=1)	1 (n=1)	1.33 (n=3)				1.33 (n=3)	2 (n=2)
Vehicle 5	3 (n=1)	3 (n=2)	3 (n=1)	3 (n=1)			1 (n=1)	2 (n=2)	
Vehicle 6	3 (n=2)			3 (n=2)	3 (n=1)	3 (n=1)			
Vehicle 7	3 (n=1)	2.5 (n=4)	3 (n=1)	2 (n=2)				2.33 (n=3)	