Individual Vehicle Differences Dominate Variation in ADAS Takeover Alert Behavior

Benjamin Bauchwitz
Department of Computer Science
Duke University, Durham, NC, 27708
Email: benjamin.bauchwitz@duke.edu

Dr. Mary Cummings, Ph.D.
Department of Electrical and Computer Engineering
Duke University, Durham, NC, 27708
Email: m.cummings@duke.edu

Word Count: 5,375 words + 1 table (250 words per table) = 5,625 words

Submitted [October 25, 2020]
ABSTRACT

Disengagement of advanced driver-assist systems (ADAS) in partially-automated vehicles where drivers may not be paying attention has significant safety implications and requires rigorous testing to ensure vehicles can safely transfer control to drivers. Existing testing and regulation on this topic are minimal, and do not address the role of difficult-to-measure but potentially important variables relating to individual vehicle differences or variations in environmental conditions. This study assessed variability in takeover alerting performance in three Model 3 Tesla vehicles in a demanding driving setting on a closed track. Results revealed that once an alerting sequence began, the ADAS systems performed as designed, but the vehicles were highly variable in terms of initiating the first alert after the loss of lane marking tracking. In addition, brightness and sun angle were found to potentially have a role in alert initiation, but dashboard-mounted cameras may be insufficient for measuring variations in atmospheric brightness. We recommend future work studying brightness, as well as more comprehensive sampling guidelines for vehicle test protocols to ensure individual differences are adequately considered.

Keywords: ADAS, driving assist, certification, individual differences, luminosity, takeover, disengagement
INTRODUCTION

At least 10% of new cars sold in the US include features enabling partial automation (1), defined as Level II Autonomy in SAE standard J3016 (2). Such vehicles include an advanced driver-assist system (ADAS) which can simultaneously perform automated steering and acceleration. However, these systems require human drivers to be alert and available at all times in case they need to take over. Events requiring the driver to suddenly takeover from the automated system are of critical importance in ADAS-equipped vehicles as changes in human attention and behavior with high levels of automation make this junction particularly dangerous (3–5). Therefore, Level II autonomous vehicles typically employ some type of driver monitoring system for assessing and interacting with drivers during these types of events. Currently, there are no US regulations addressing how driver monitoring should be implemented or the performance standards such systems should meet.

Formal testing of ADAS systems in general is limited. The National Highway Traffic Safety Administration’s (NHTSA) New Car Assessment Program (NCAP) does not address driver-assist features, limiting its assessment to collision and rollover survival (6). The Insurance Institute for Highway Safety (IIHS) covers some ADAS features such as pedestrian detection and automated emergency braking (AEB), but does not address driver monitoring (7). The European NCAP has similar scope, but has announced that it will begin assessing driver monitoring in the 2022 revision of its protocols (8). The Korean Ministry of Land, Infrastructure, and Transport (MOLIT) is the only major authority regulating driver monitoring, which it has done as of June 2020 (9), and addresses specific guidance regarding how the system should handle engagement and disengagement of automated features.

Even in the few cases where guidance is provided on how driver monitoring and takeovers should be evaluated, certain aspects of testing are still highly ambiguous. In particular, there is little to no specification as to how individual differences in vehicles or variations in operating environment should be considered in the test procedure. None of the NHTSA, IIHS, NCAP, or MOLIT test protocols address how vehicles should be sampled to ensure the test results are robust to differences in vehicle trim, configuration, or wear-and-tear. Likewise, these test processes either take place in sterile laboratory settings or make assumptions about the generalizability of the results to different atmospheric conditions or road environments. They may especially underestimate the effects that subtle changes in atmospheric conditions have on vehicle performance.

The goal of this research was to evaluate the impact of these under-studied test parameters on the behavior of a driver monitoring system for an ADAS-equipped vehicle faced with an extreme maneuvering environment. Given that Teslas have ADAS systems that can be used on interstates, divided highways, and urban and rural roads and thus, can face a range of potentially demanding environments, we elected to use Teslas as our test platform. In the next sections, we will describe how driver monitoring works in the Tesla Model 3 and the experimental setup, and detail the results and analysis.

METHODS

The goal of this experiment was to determine if there is significant within- and between-vehicle variability in the type and timing of feedback presented to the driver when an ADAS-enabled vehicle can no longer adequately track lane markings. Given that computer vision is based on probabilistic reasoning with a potential large performance range, our hypothesis was
that there would be variation in the timing and duration of events in this sequence both between vehicles as well as for successive tests of the same vehicle.

Three 2018 Tesla Model 3s from the Triangle metropolitan area of North Carolina were randomly selected for study using a car sharing service over a one-week period during March 2020. Tests were conducted at the North Carolina Center for Automotive Research (NCCAR), a closed test track facility. The NCCAR test track is a two-mile long, 40-foot-wide paved loop with a mix of straightaways and curves of a widely varying range of angles. All tests were conducted during daylight, between 1:00pm and 5:30pm, under similar visibility, cloud cover, wind, air pressure, and humidity. The cars had different software versions, including v10.2 (2020.4.1 4a4ad401858f), v10.2 (2020.12 4fbcc4b942a8), and v10.2 (2020.8.1 ae1963092ff8).

Autopilot settings were consistent across trials, but some vehicle settings varied since the cars were randomly chosen and could not be changed. These included the acceleration style, the appearance of the navigation screen, and whether the built-in dashboard cam was setup to save video. None of these factors were confounds for this test. The same person drove the vehicle for all tests. Prior to each trial, the vehicle was placed in park, with the driver exiting and using the key card to lock and deactivate the vehicle before entering the car to begin a test.

**Tesla Alerting for Driver Takeover**

When Tesla Autopilot can no longer confidently track lane markings on the road, it requests that the driver take control of the car. Teslas recognize that a driver has taken control through a torque monitoring system on the steering wheel that measures how forcefully the steering wheel has been rotated in an attempt to infer whether the driver has deliberately manipulated it.

Requests in the Tesla Model 3 are presented on a 15-inch, 1920 x 1080 pixel display centered between the driver and passenger seats. This display also displays all information about the car, including speed and rpm, as well as a map. When a Tesla requests that the driver takeover control, these requests manifest as a sequence of three distinct alert phases. First, the console displays an icon and text requesting the driver to grasp the steering wheel (Figure 1a), and is accompanied by a quick pair of beeps.

**Figure 1: Progression of alerts in the takeover sequence.**

Left: The first alert includes a black bubble with an icon and the message “Apply slight turning force to steering wheel.”

Center: The second alert has the same icon and text, but introduces a flashing blue light at the top of the screen.

Right: The third alert displays a large red icon with hands on the steering wheel and includes a red bubble with the message “Autosteer unavailable for the rest of the drive.”
If the driver does not grasp the steering wheel after the first request, after approximately 10 seconds the top left of the Autopilot display on the console flashes blue with progressively increasing intensity (Figure 1b), culminating in two pairs of beeps. If the second request for driver control is ignored, after approximately 5 seconds the small icon and text are replaced with a large red icon and red highlighted text (Figure 1c), accompanied by three pairs of beeps. The hazard lights then activate and the car will slow to a stop. The combined total alert duration from the first to third alert is 15 seconds (10). The car is intended to begin the alert sequence within 400 elapsed milliseconds of the loss of lane markings under these circumstances (10).

Test Protocol

We created a demanding driving scenario on the NCCAR test track in which the Teslas encountered the disappearance of lane markings and subsequent extreme curves. This setup was intended to degrade the vehicle’s model of the roadway to force initiation of the driver takeover sequence. To test the driver monitoring system, the driver did not respond to the alerts to take over control, but was ready to do so at any point a trial became unsafe. This test utilized a section of track at NCCAR including an approximately 600-foot straightaway followed by a series of sharp curves (Figure 2). Prior to the data collection, a 330-foot section of straightaway was marked with highway-style white lane lines (10 feet long by 6 inches wide with 30 feet of longitudinal spacing between) to form three lanes 13 feet wide.

The vehicle began between a pair of traffic cones at the position marked “start” in Figure 2, approximately 250 feet before the painted section of the track. Starting from its inactive, parked state, the vehicle was manually driven towards the painted lane lines in the inner most lane. After accelerating to 35 miles per hour, the car was immediately placed in Adaptive Cruise Control to fix the speed. After passing the cones at the end of the fourth lane marking, Autopilot was activated and the painted lines ended after the initial straightaway in order to cause loss of lane tracking by the computer vision system. Each car then experienced two sharp curves, approximately 120 and 190 degrees. The car was allowed to drive autonomously until the system reached the third alert without driver response, which resulted in Autopilot shutting down the vehicle. Each vehicle was subjected to 10 repetitions of this test, but the order of the tests was randomized across the three cars.

Data Collection and Analysis

Video data were collected using three GoPro Hero 7 Black cameras synchronized with SyncBac Pro devices and mounted at fixed positions in the vehicle interior. The first camera was mounted on the dashboard facing forward to monitor the roadway. The second camera was hung upside down from the sunroof facing forward to monitor the center console. The third camera was mounted behind the steering wheel facing backward to monitor the driver. All cameras were set to 1440 pixels per inch resolution, 25 frames per second, wide field of view, and automatic stabilization, with protune off. Protune off meant that certain camera properties, such as the exposure, adjusted automatically in real-time based on the light being received by the camera.
Videos were manually scored to identify the first frame at which events of interest occurred and SyncBac timecode annotations were used to link the corresponding frames taken from different cameras. The events that were scored were the timing of each of the three alert stages and the time at which the vehicle first entered the initial curved portion of the track.

Given that brightness could influence camera perception, information about scene brightness was extracted from video frames by converting the videos from RGB to YCrCb color space using the Python OpenCV image processing toolkit (11). The brightness emanating from a particular region in a visual scene is defined as that region’s luminance, is measured in candela per square meter, and can be obtained from a weighted sum of the red, green, and blue light in that region (12). However, as a result of humans’ limited capability to perceive differences in the...
intensity of light, light conveyed in images is compressed by several orders of magnitude using a process called *gamma compression*, where some information is lost (13). The “Y” channel in a YCrCb image encodes the image’s *Luma*, which is as close as an approximation of true luminance as can be obtained from gamma compressed digital imagery (14). Luma is a unitless index taking a value between 0 and 255. Other color models that encode light intensity, such as HSV, HSI, or HSL, do not match the perceptual qualities of brightness as well for certain colors (15), making YCrCb superior for extracting brightness from images.

We restricted our analysis of brightness to only the top 380 rows of pixels in the images from the forward-facing roadway camera, which corresponded to regions of the scene that only contained the sky. This was done to reduce the influence of sudden changes of color, such as variations in the treeline or the appearance of distant objects. A number of studies have shown strong results in image processing tasks using luminance-based color models, including for tasks related to perceiving pixels in the sky (16–20).

**RESULTS**

While the primary purpose of this study was to determine whether there were any variations between and within cars in terms of how the stages of takeover alerts were presented to the driver, it is worth noting how well the cars performed in general in terms of controllability.

Despite the challenging course and the driver’s intentional ignoring of the takeover requests, Autopilot successfully maintained control through the test durations and brought the cars to a safe stop in all 30 trials.

The next sections will discuss the variations in the timing of the takeover alert sequence as well as the location at which it was presented.

**Timing of Alerts**

In order to determine the consistency of the three stages of alerts within and between cars, two measurements were used to assess the length of the alert sequence stages: the length of time between the first and second alerts and the length of time between the second and third alerts. Because the third alert coincided with the end of autonomous driving, no time measurements were taken past this point. The total duration of the alert sequence was also measured, which was the length of time between the first and third alerts.

Variation in the time period between the first and second alert was minimal and corresponded with the 10s specification provided by Tesla, with a mean of 10s and a standard deviation of approximately 0.1s across the entire dataset (Figure 3a). A repeated measures ANOVA did not reveal any significant differences between cars or within cars across the different trials.

Variation in the duration of the second alert stage was also small, although car #2 skipped this second alerting stage (Figure 3b) in one trial. Regardless of this outlier, all three vehicles had a median value of 5.0s, again matching the 5s Tesla specification. A repeated measures ANOVA did not reveal any significant differences between cars or within cars across the different trials.

Variation in the overall duration of the alert sequence was also small, again with the exception of a single outlier in car #2 (Figure 3c). Again, the data for all three cars closely matched the Tesla specification with a median duration of 15s across the dataset. A repeated measures ANOVA did not reveal any significant differences between cars or within cars across the different trials.
Location of Alerts

While we did not detect any meaningful variation in the timing of the takeover alert sequence once it was initiated, we did detect significant variation in where the first alert was initiated. The first alert location was estimated by comparing the time the first alert occurred with the time the vehicle first entered the curved portion of the track (marked by an easily identifiable set of traffic cones, Fig. 2). Approximate distance was computed from this time interval by multiplying the vehicle’s reported speed (fixed at 35 mph for the test) over the time interval traveled with the position confirmed through a manual inspection of the video.

Figure 4 shows the approximate position of each car when the alert sequence was first initiated for every trial. Each trial is represented with a car-specific mark, where car #1 is depicted by a red X, car #2 a yellow circle, and car #3 a blue plus sign. Each mark illustrates the approximate longitudinal position along the track, with lateral variation included solely to improve readability. It is important to remember that all 30 trials began in the same place and at the same speed, with Autopilot engaged.

It is clear from Fig. 4 that there were three distinct areas where the first alert occurred, with separation between cluster centers. An iterative K-means analyses using 1-10 clusters and assessed with the Gap statistic yielded K = 3 statistical clusters (Figure 5). The first cluster (Zone 1) was centered at 121 ft after the beginning of the curve, with a range of 43 to 172 ft; the second cluster (Zone 2) was centered at 637 ft with a range of 609 to 768 ft; and the third cluster (Zone 3) was centered at 1248 ft with a range of 1240 to 1255 ft. The distribution of these distances for each car are shown in Figure 6a.

Clusters were not equally represented across cars; Car #1 and Car #3 experienced alert initiation points in Zones 2 and 3, while Car #2 experienced alert initiation points in Zones 1 and 3 (Table 1, Figure 6b). A chi-squared independence test revealed a significant difference in the counts between cars ($\chi^2 = 22.4$, $p = 0.0002$).
Figure 4: Approximate location of the first alert for each trial
Figure 5: Optimal number of clusters in data for distance from curve to alert sequence initiation based on K-means clustering with the Gap statistic

<table>
<thead>
<tr>
<th>Alert location</th>
<th>Car #1</th>
<th>Car #2</th>
<th>Car #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone 1</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Zone 2</td>
<td>9</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Zone 3</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

a. Distance traveled by each car from first curve until first takeover alert is presented

b. Number of alerts at each zone by car

Figure 6: Position on the track for first alert in the takeover sequence. Left shows distribution of distance traveled by each car while right shows the counts of cars in each zone.

Investigating the Different Initial Alert Locations

Given that there were clear differences in where the cars experienced the first alert due to the car’s loss of lane marker tracking (Fig. 5), we examined which independent variables could predict the location of the first onset of an alert in order to determine possible sources of variation.

To this end, we first examined whether there was any difference between the brightness levels perceived by cars with and without alerts in Zone 1. To do this, we computed the luma for each car as outlined in the Methods section for the 150-ft segment of track immediately preceding Zone 1, a distance of roughly 3s prior to an alert. An analysis of covariance model was developed with the luma as the dependent variable and alert zone as the main factor. Sun angle (azimuth), based on the date and time each trial was conducted, was included as a covariate because it has been cited as a potential source of problems for Autopilot (21).

The results were marginally significant (F(1,27) = 3.87, p = .06 with an alpha of .05), meaning that the trials with alerts in Zone 1 experienced different brightness levels than those without alerts. The sun angle covariate was marginally significant (F(1,27) = 3.39, p = .08). Car 2, the only car to experience the initial alert in Zone 1, had a mean luma of 178 and the rest were at a mean of 163. The mean sun angle for car #2 in Zone 1 was 189 degrees and 212 degrees for the no alert group in Zone 1. Thus, one possible influencing factor is not only how much brightness is experienced when lane lines are lost, but also what the sun angle is.

A backwards logistic regression (LR) model was used to predict whether the alert occurred in Zones 1 and 2 as a function of sun angle and brightness in the 3s prior to each alert.
In this model, luma was also marginally significant (p=.06, B = -.176) and sun angle was not (model accuracy = 85%). Because LR models produce regression coefficients for each feature that are log odds, taking the exponential of the coefficient weights estimates the expected change in the log odds of the target variable per unit increase in the corresponding predictor variable, holding the other predictor variables constant. This means that as luma increased by one unit, there was a 12% increase in the likelihood the car would experience a Zone 1 initial alert.

DISCUSSION
We found that the presentation of the takeover alert sequences, once initiated, was consistent between and within vehicles. The takeover alert sequence successfully activated in all 10 trials of all 3 vehicles studied, and these takeover alerts were nearly identical across every trial and vehicle. In all 10 trials for cars #1 and #3, the alert sequence contained all 3 of the expected stages (Figure 1) without any variation in the types of sounds or icons that were presented. For car #2, 9 of 10 takeover alerts contained all the expected stages, but in one trial, the 2nd stage of the alert was skipped.

While takeover alerting was consistent once initiated, the occurrence of the alerts was neither consistent within a single car nor across all three cars. This is especially noteworthy since all trials were conducted on the same section of track, under laboratory conditions, at the same time of day and under approximately equivalent weather conditions. Distances between the end of marked lanes and the initial alert could be as short as 43 feet (13 meters) and as late as 1255 feet (383 meters). Moreover, in 30% of trials, cars travelled 26s beyond the point of loss of lane markings before warning the driver. This is particularly important in light of the Tesla Model 3 Mountain View crash in 2018 which killed the driver, where loss and/or degradation of the lane marking was discussed as a potential failure mode (10).

The significant degree of variability across and within cars is puzzling, as it suggests unknown possible triggers for the alerts. The loss of lane markings in the straightaway prior to Zone 1 is likely the cause for half of car #2’s initial alerts, but may not be the cause of its remaining alerts in Zone 3. It is not clear if the alerts in Zone 2 are delayed responses from the loss of lane markings in the straightaway or from some new stimulus. It is also unclear what led to Zone 3 alerts, which occurred 26s after the loss of lane markings. These may be explained by some interaction between the computer vision system and possible steering wheel angle saturation rather than loss of lane tracking, given the 190-degree curve, but more work remains to identify this cause.

When looking at just the Zone 1 and 2 alerts, there was some evidence that brightness and possibly sun angle could be influencing factors. Sunlight has been shown to cause anomalies in the perception systems used in these and other autonomous vehicles (22–25), so it is possible that sun angle could contribute to a camera system’s degradation. While these results suggest that there could be an important connection between brightness, sun angle and the triggering of camera vision-based alerts, more work is needed to further investigate these findings.

Individual variability within cars was high and there was no consistency in location of alerts for any of the three cars. Cars #1 and #3 experienced alerts in Zones 2 and 3, while car #2 experienced all Zone 1 alerts and 56% of Zone 3 alerts. It should be noted that car #2 had the full self-driving upgrade and the owner complained about significant Autopilot issues. This issue highlights that although the vehicles tested were all the same make, model and year, they were not identical machines. They all had slightly different trims with slight differences in the options
enabled, and because of Tesla’s over-air update policy, all were running different versions of the Autopilot software.

One limitation to our approach is that the camera used to detect the luma values automatically adjusts the exposure to equalize the appearance of images despite differences in available light. While we did detect variations in luma of up to approximately 20% across images, this may not have represented the true variation in light across the different scenes. More work is needed with light meters for better accuracy. In addition, the vulnerability of computer-vision systems engaging in lane tracking to different sun angles deserves further scrutiny, as well as the interaction between computer vision systems and steering wheel angle saturation.

CONCLUSIONS

This study addressed the behavior of Tesla vehicles during the transition from automated to manual driving when human drivers were alerted that they needed to take control after the loss of lane markings. We investigated whether there were differences in the takeover sequence between different cars as well as within cars.

This study demonstrated that after losing lane markings, once each car initiated the three-stage alerting sequence requesting the driver takeover at each stage, all cars flawlessly executed the alerts in the appropriate time intervals and stopped the car when the driver did not respond to the alerts. However, there were widely varying differences across the cars as to when each car initiated the first warning to the driver after losing the lane marking. Since vehicles were consistent on all Autopilot- and alert-related settings, it is unlikely that any other internal vehicle differences contributed to the observed effect.

In the 30 trials, the first alert requesting handover of control from the car to the driver occurred up to one-quarter mile from the loss of lane markings and could take up to 26s with no human intervention. A previous accident demonstrated that lane markings could be lost and with just 6 seconds of driver inattention, a fatality could result (10). Manufactureres of ADAS systems need to be vigilant in ensuring perception systems reliably recognize when they are operating outside of their competency boundaries and immediately notify the driver.

This study also demonstrated that there might be important relationships between external brightness and sun angle and when a car may alert drivers to a loss of lane markings. This speaks to the larger issue of possible sources of brittleness for computer-vision systems and the need to determine how changes in the environment can degrade outcomes. To this end, future research should include dedicated measurement devices for assessing brightness, as cameras may not have adequate perceptual power to discriminate between different light intensities. Also, more dedicated, realistic testing is needed for both the cameras and the integrated systems to ensure that such wide variability as seen in this study is understood and mitigated.

More broadly, this research highlights the potential for significant variability between individual vehicles of the same make, model, and year. Current practices for certifying vehicle technology generally rely on testing only a single car, and generally limit testing to major design changes, rather than testing all unique variants. For complicated components like ADAS systems, this practice may fail to identify the full range of possible behavior in vehicles as it may not detect flaws that are either latent or present inconsistently. Moreover, incorrect conclusions may be drawn from a single vehicle that performs differently from the broader population of fielded vehicles. The important role of individual vehicle differences should be considered as ADAS testing standards continue to develop.
ACKNOWLEDGMENTS
This research was funded by a US Department of Transportation’s University Transportation Center grant through the University of North Carolina’s Collaborative Sciences Center for Road Safety. We were assisted by Sam True, the director of the North Carolina Center for Automotive Research, and Matthew Seong, Kausthub Ramachandran, and Vishwa Alaparthy in the collection of the data.

AUTHOR CONTRIBUTIONS
The authors confirm contribution to the paper as follows: study conception and design: B. Bauchwitz, M. Cummings; data collection: B. Bauchwitz; analysis and interpretation of results: B. Bauchwitz, M. Cummings; draft manuscript preparation: B. Bauchwitz, M. Cummings. All authors reviewed the results and approved the final version of the manuscript.
REFERENCES


