Safe and Sustainable Fleet Management with Data Analytics and Training

A Technical Report submitted to the Department of Engineering Systems and Environment

Present to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Josh Kim

Spring, 2021

Technical Project Team Members
Thomas Gresham
James McDonald
Nick Scoggins

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signature: ***Josh Kim*** Date: ***05/05/2021***

Approved ____ Date ___________
Brian Park, Department of Engineering Systems and Environment

Safe and Sustainable Fleet Management with Data Analytics and Training

Thomas R. Gresham
B.S. Systems Engineering
University of Virginia
Charlottesville, VA
trg3ua@virginia.edu

Joshua Kim
B.S. Systems Engineering
University of Virginia
Charlottesville, VA
jk3ct@virginia.edu

James McDonald
B.S. Civil Engineering
University of Virginia
Charlottesville, VA
jm7sh@virginia.edu

Nick Scoggins
B.S. Systems Engineering
University of Virginia
Charlottesville, VA
nss5ys@virginia.edu

Moeen Mostafavi
Engineering Systems and Environment
University of Virginia
Charlottesville, VA
mm4ff@virginia.edu

B. Brian Park
Engineering Systems and Environment
University of Virginia
Charlottesville, VA
bp6v@virginia.edu

Michael D. Porter
Engineering Systems and Environment
School of Data Science
University of Virginia
Charlottesville, VA
mdp2u@virginia.edu

Michael E. Duffy
Transportation Operations and Fleet Manager
University of Virginia
Facilities Management
Charlottesville, VA
med7p@virginia.edu

Sandra A. Smith
Manager, Quality Assurance & Staff Development
University of Virginia
Facilities Management
Charlottesville, VA
sas7rs@virginia.edu

***Abstract* - In scores of vehicle fleets, telematic tracking systems provide fleet managers with information regarding energy consumption, the obedience of safety regulations and driver performance. For a University's Facilities Management (FM) Fleet to take the next steps towards an elevated Sustainable Fleet accreditation and overall team performance, the management has recognized the importance of effective energy and safety tracking methods combined with data analytics and a comprehensive systems analysis in order to aid the reinforcement, training and maintenance of safe and sustainable driving practices by fleet drivers. This paper outlines the design of a unique safety and eco-driving training program that will prompt University FM drivers to reflect on, educate and develop mindful driving habits which reduce environmental impact, cost and risk. We analyzed historical driver behavior data, including idling time, harsh acceleration, crash incident details, resulting in the identification of risk factors and areas for significant improvement. Educational aspects of the training program were influenced by focus groups, interviews administered with industry experts, and professional fleet training modules. The efficacy of the customized training program has been assessed through a statistical evaluation of telematic data collected before and after the training program was delivered. The results indicate that agency-specific mindful driving training was statistically significant in improving five of the six behavioral metrics measured - idling time, seat belt usage, speeding, hard acceleration, and hard braking - compared to the control group.**

Keywords - Data Analytics, Mind-Driving, Training Program

I. INTRODUCTION

Fleet managers are responsible for ensuring departmental adherence to regulations, tracking and maintaining fleet vehicles, and optimizing organizational costs. With the duty to continually improve safety and sustainability metrics of the fleet, eco-driving has emerged as a novel approach. Eco-driving can be classified into three particular groups: strategic decisions (vehicle selection and maintenance), tactical decisions (route planning and weight) and operational decisions (driving style) [1]. Historically, tactical and strategic decision-making has been the prime function of fleet managers, but with the rise of telematics, in-vehicle sensors that capture driving behavioral metrics, managers are now capable of monitoring driver performance regularly. Eco-driving training has proven to be effective in reducing fuel consumption by up to 20% while also decreasing crash risk [2]. However, given the range of contextual differences across past eco-driving training trials (vehicle types, country of origin, driving routes, etc.), there is no standardized methodology for a successful educational program. Additionally, each fleet driver has their own distinctive driving style. Obtaining a baseline on the driving behaviors contributing to the greatest fuel inefficiencies and compliance concerns is important to support feedback on driving. Yet to date, no training program has been driven by agency-specific data, nor focused on the promotion of both safety and sustainability (mindful) best practices.

A University's Facilities Management (FM) Fleet, a team composed of roughly 260 vehicles ranging from lightweight vans to heavy duty truck cabs, has furthered sustainability goals' through the replacement of less-efficient diesel vehicles with electric and hybrid vehicles. To continue to reduce the fleet's carbon footprint, enhance the current sustainability-minded accreditation and promote basic compliance-related behaviors, we propose an interactive data-driven training program which provides basic information about safe and sustainable driving coupled with data regarding the fleet's holistic driver performance historically. This mindful-driver training targeted incident counts on harsh braking, speeding, acceleration, cornering, idling, and seat belt violations. Accordingly, this study was conducted on the University fleet of vehicles and organized in three phases as follows: (1) baseline vehicle data collection/analysis (pre-mindful-driver training); (2) training program pilot development and implementation (mindful-driver training); (3) post-training vehicle data collection/analysis (post-mindful-driver training).

II. Background

Transportation is central to sustainability and safety risks in the United States of America (USA). In 2018, the combustion of fossil fuels for cars, trucks, ships, trains, and planes accounted for 28.2% of greenhouse gas emissions, the greatest proportion of emissions in the USA [3]. In urban environments, commercial vehicle fleets account for nearly 20% of total mileage [4]. In order to address climate change concerns, government and commercial fleet organizations have primarily turned to the development of more fuel efficient technologies, such as electric vehicles and automated engine control in-vehicle mechanisms. Due to resource constraints and technological challenges, the mainstream adoption of these fuel saving innovations is more likely a long-term solution. Within the short-term, eco-driving has emerged as a promising method for reducing overall fuel consumption-related emissions by influencing changes in driver behavior. Eco-driving techniques contribute not only to a greener environment, but also to increased safety. The combination of smooth driving and anticipating traffic conditions, which are essential to improving fuel economy, make drivers less likely to be in an accident [5]. This is a significant benefit given that the leading type of fatal work-related event in the USA are car crashes [6] and driver behavior has been shown to be a major causing factor [7]. Thus, extensive opportunities exist for commercial vehicle fleets to provide mindful driver behavioral intervention training due to the combination of safety and environmental benefits at low cost. By analyzing telematic data, significant benefits and areas for improvement of training can be identified by fleet managers to inform subsequent decision-making. Currently, fleet training programs are generic meaning they do address the specific areas a fleets' drivers need to improve in. We noticed this while reviewing industry examples such as the NTSI, National Traffic Safety Institute, SAFER Driver Fleet and Distracted Driver Avoidance Course [8]. Both of these courses cover a lot of information but are not targeted towards the relative deficiencies of the drivers taking the course. This leads to drivers not focusing on what they need to improve on and reduces the effectiveness of the training. Additionally, these courses did not connect safety to eco-driving, which combine to form mindful driving.

III. Methodology

Our goal was to design an agency-specific mindful driving training program to reduce incident counts of compliance and fuel consumption-related fleet behavioral metrics. After gaining an understanding of the fleet improvement needs through baseline data exploration, stakeholder engagement and requirements gathering, we sought to develop an hour-long training program to address and ameliorate these specific poor driving habits. We implemented this training in a pilot study on a select group of historically poor performing drivers, and with the help of a trained FM development professional. Finally, once the pilot study was completed, we analyzed the effectiveness of the training by comparing the statistical significance of average incident rates of the training group versus a control group of drivers.

A. Baseline Vehicle Data Collection/Analysis

Comprehensive data collection and analysis was based on one main dataset: driver behavior incident count data. The big data was provided via IoT in-vehicle sensors tracking each fleet vehicle and transmitted to a central cloud database. Upon initial analysis of the various structured and unstructured formats of the big data, our team decided to focus on structured big data - weekly aggregates of driver performance, commonly referred to as "scorecards." This decision was made on the basis of automatic summarization. Our team considered utilizing unstructured data containing daily raw incident counts per vehicle, however this format recorded vehicle status every few seconds, leading to data overlap and data transformation complexity. It was deemed that the weekly scorecards held all the necessary metrics for evaluation and further efforts to manipulate the raw unstructured data would ultimately lead to a similar structure as the scorecards.

The weekly scorecards contained the following data metrics - instances of 1) hard acceleration, 2) hard braking, 3) hard cornering, 4) speeding, and 5) seat belt violation while driving. Additionally, the scorecards contained information on 6) total time spent idling in minutes. For the purposes of this study, an incident for each metric was specified when exceeding the following defined in Table 1.

Table 1. Metrics and definitions of how incidents were recorded

	Metric	Metric Threshold Definition
1	Hard Acceleration	Force of 0.38 G in the backwards direction
2	Hard Braking	Force of -0.56 G in the forward direction
3	Hard Cornering	Combined acceleration and natural force of 0.41 G
4	Speeding	5 mph above posted speed limit
5	Seat Belt	Seat belt violation while car is moving over 1 mph with repeated incidents recorded every 15 seconds

Hard acceleration, braking, and cornering incident counts have the potential to indicate a vehicle's likelihood to engage in risky behaviors resulting in crashes and/or fatalities. Hard acceleration is characteristic of a driver attempting to weave through traffic, being inattentive at a traffic light, or needing to make up lost time for task completion. Hard braking takes place when a driver neglects proper following distance or is inattentive to current environmental conditions such as the presence of pedestrians. Speeding, or rather, exceeding the posted speed limit may result in unnecessary accidents. Negligence of seat belt usage impacts fatality rates during accidents. Speeding and seat belt usage also carry legal consequences which influenced the necessity to include these metrics. Furthermore, all of these metrics are factors relevant to safe driving.

With respect to sustainable driving, unnecessary idling time and hard acceleration were examined as two key metrics which heavily influence carbon emissions [5]. Certain tasks require a level of idling, however, our team identified a need to challenge FM drivers to minimize time in non-essential idling situations. Hard acceleration affects fuel consumption as an increase in force to move a non-stationary vehicle quickly requires more energy.

With an understanding of how each of the selected metrics in data could be used to influence behavior, we then joined the weekly scorecards data with shop level data to allow analysis at the vehicle or shop level by week. Once the structural components of the data were cleaned, we trimmed the data set to include only weekly scorecards over a one-year period from September 2019 to September 2020. This selected data was further analyzed to provide an overarching view of fleet tendencies and inefficiencies.

B. Training Program Development and Deployment

We developed a training program geared towards this agency's fleet team to increase safety and sustainability. After multiple discussions with FM, our team decided to conduct a pre-training discussion and develop training modules geared around Safety and Eco-Driving.

Throughout the development process our team identified three key elements which we believed to be highly influential in the success of our training program. Those elements were transparency, inclusion, and motivation.

With regard to transparency, we believed the drivers should know that the tracking devices are installed on their vehicles and monitoring their performance at all times. We also thought drivers would be more concerned with what data is being collected, and how it's being used to alter their behavior. Therefore, we aimed to provide drivers with real-world applications of our training, rather than inundating them with technical definitions of the performance metrics. For example, in our training program rather than giving a technical definition for harsh braking and acceleration incidents, we calculated the time allowed for a vehicle of a certain size and weight to accelerate (brake) from 0 (25) miles per hour to 25 (0) miles per hour, without the monitor recording an incident.

In the development of our training program, surveys played a key role in indirectly including drivers in the development. Hearing from the drivers themselves about the ways they learn best allowed us to include elements in the training that would be relevant to every type of learner in the audience. We also allowed some drivers to sit in on meetings where we presented selections from the training throughout its development. Gathering first-hand feedback from drivers throughout several iterations of our training program gave them a sense of ownership over their data, while also revealing to our team small tweaks to the program that proved to be highly influential in its effectiveness.

The high amount of motivation on behalf of FM management was clear through their continued support and encouragement of our team's efforts. This was crucial as an organization cannot expect buy-in from its employees if its leaders don't do so themselves. The public support of our project by leaders within FM played a vital role in obtaining the level of commitment from drivers that we did. Additionally, we found consulting with an expert training facilitator to be invaluable. Sandra Smith, the expert, helped our team craft a training program that we believe was interactive, engaging, and applicable to real world situations. Experts such as Sandra can have a huge influence on obtaining the necessary buy-in from drivers.

Module Development

The Safety module was designed to tackle metrics such as hard acceleration, braking, and cornering, along with seatbelt usage and speeding. We input fleet crash data and information provided by the Fleet Manager to help us identify the causes and costs of crashes within the fleet, and further used this agency-specific data paired with general educational knowledge. For example, as seen below, by coupling the most common cause of crashes specific to the agency along with both a question asking agency participants what some vehicle distractions are, how to combat distractions and providing data regarding driver distractions in general, participants are engaged with the training and retain knowledge pertaining to themselves.

Fig 1. Sample Safety Module Slide

The Eco-Driving module focused on identifying and improving sustainable driving practices within the fleet. As a result, our team focused on metrics such as idling time, hard acceleration, and extreme speeding. A majority of the module focused on idling specifically because of the significant impact it has on greenhouse gas emissions.

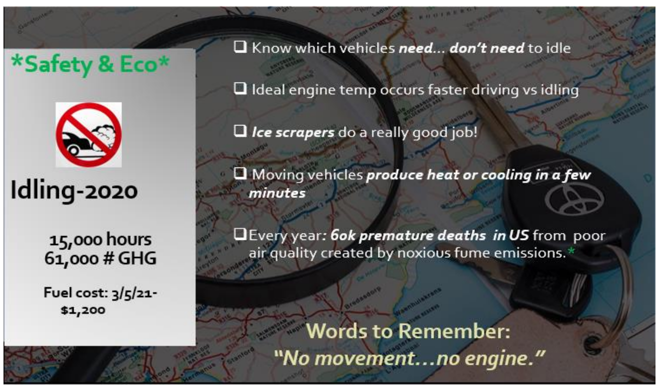

Fig 2. Sample Eco Module Slide

Vehicle Selection for Training Program

Given the size of the fleet and the complexity of scheduling, 4 shops consisting of 2-3 vehicles were selected to participate in the pilot program. Each vehicle is driven by multiple drivers and due to privacy concerns, drivers were not matched to specific vehicles.

Shops considered for this program consisted of having between 2-4 vehicles, have vehicles driving for at least 60% of the time range the data was collected, and consisted of vehicles that had significant incident counts spread across all metrics. Our team provided FM with the 4 shops for the training pilot and the supervisors for these shops were notified that they were selected to help build a training program. The drivers for these shops were not notified about their driving scores. A total of 10 vehicles were analyzed from these groups.

To create a control group, our team found comparable vehicles for each of the 10-vehicles based on each individual metric. This process compared mean incident counts along with the variance, and also looked at the average distance driven. Table 2 further clarifies the vehicle comparison process.

Table 2. Example comparison between 1 vehicle from control group and 1 vehicle from training group

Vehicle group	Incident Average	Incident Variance	Total distance
Control	266.5	23889	78.44
Training	240	24161	60.8

Training Program Deployment

As a result of the ongoing pandemic, the training session was unable to be held in-person and was moved to an interactive online video conference session as portrayed below. The training was held on 3/24/21.

Fig 3. Training program being delivered via online video conference

C. Post-Training Vehicle Data Collection/Analysis

During the training, the drivers were asked a series of questions. These questions served to engage the drivers and their resulting answers collected post-training demonstrated that the majority of drivers had a fair understanding for the components of distracted, safety and eco-friendly driving habits, even though participants didn't implement them to the fullest extent before training

To evaluate and compare pre- and post-pilot driving results, pre-pilot data was collected between 2/24-3/23 and post-pilot data was collected between 3/25-4/21. All metrics were converted into incident rates by dividing the sum of total incidents as shown in the equation below.

For each vehicle metric, the data was normalized as follows:

$$Incident\ Rate\ = \frac{Sum(Incident\ Counts)}{Sum(Miles\ Driven)}$$

IV. DATA ANALYSIS AND RESULTS

Data Analysis

Our team looked at 4-weeks' worth of driver scorecards to gauge the effectiveness of the created training program. Hypothesis tests were run on the training and control groups to determine if there was a statistical difference between the incident rates per mile before and after the training. A threshold of α = 0.05 was used.

$$H_O: Before\ Rate\ -\ After\ Rate\ =\ 0$$

$$H_A: Before\ Rate\ -\ After\ Rate\ > 0$$

After checking the distribution of the differences for each metric, our team determined that the data was non-normal based on the p-values of the Shapiro-Wilk normality test and opted to conduct Wilcoxon-Signed-Rank tests instead of t-tests. The Shapiro-Wilk normality test established that the scale of data for hard acceleration, hard braking and hard cornering must be normalized based on the p-values on the differences of incident rates. The p-values from conducting a Shapiro-Wilk normality test were observed as: hard acceleration (2.7E-6), hard braking (8.1E-4) and hard cornering (4.6E-6).

The Shapiro-Wilk normality hypothesis test is as follows:

$$H_O: Distribution\ of\ Data\ is\ Normal$$

$$H_A: Distribution\ of\ Data\ is\ Non-Normal$$

Figures 4 and 5 show the changes in rates across all metrics before and after the training program period. Vehicles in the control group experienced both increases and decreases in rates while the training group had a downwards trend for most of the vehicles. One vehicle in the training group, however, has a driver nearing retirement which accounts for the singular upwards trend across all respective metrics.

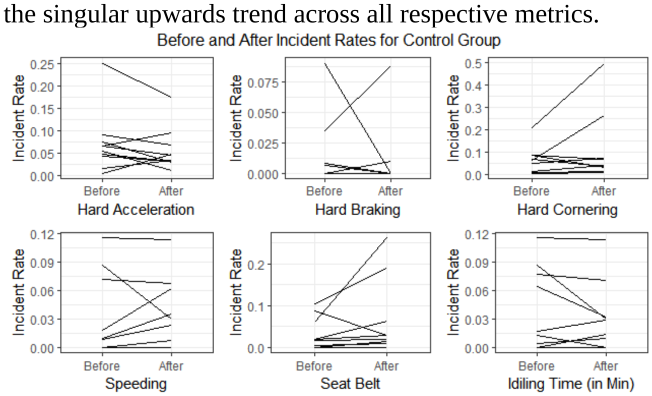

Fig 4. Before and After Plot of Control Group

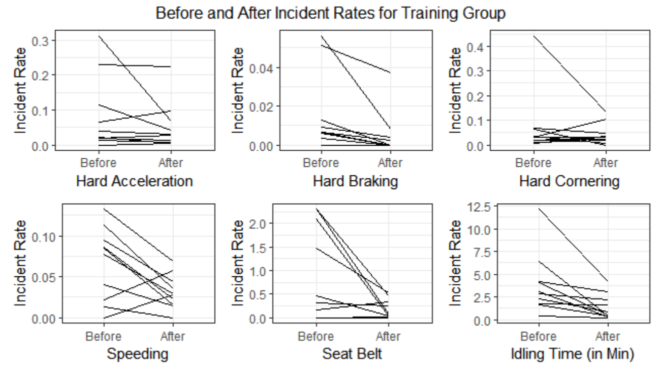

Fig 5. Before and After Plot of Training Group

Based on the results from Table 5, our team was able to identify that the training program was effective in decreasing incident rates across most metrics while the control group was not significant across any metric used. More specifically, with the exception of one vehicle, all vehicle operators supplied with driving reduced their incident rates. As a result, we can identify that our training program was effective in addressing both behavioral and compliance elements within our training group. It can be argued that complexity surrounding what cornering is and a lack of focus on this metric in the training program led to less noticeable change by trained driver participants.

Table 3. P-values from conducting Wilcoxon signed-rank tests for each metric

Metric	P-Value Control	P-Value Train
Hard Acceleration	0.138	0.161
Hard Braking	0.500	0.007
Hard Cornering	0.813	0.423
Speeding	0.777	0.019
Seat Belt	0.221	0.001
Idling	0.960	0.029

With regards to eco-driving, the training group experienced an average decrease of 1.34 idling minutes per mile driven compared to the control group which logged a difference of 0.01 idling minutes per mile driven. These results help show that this agency-specific training program has the ability to address both driver behavior and general compliance.

V. Conclusions, Limitations and Future Work

Through our literature review, it was found that no professional fleet-driver training program on the market is data-driven agency-specific, and most trials, if not all, solely study the consequent effect of training on fuel consumption. Our pilot program was developed to be unlike any previouslymade course. Our training delivered in the two-week pilot analysis showed a significant reduction in hard acceleration, hard braking, speeding, seat belt usage and idling. In the four-week pilot analysis, the effects on hard acceleration incident counts diminished to be insignificant between the control and training group. Given the findings in the baseline analysis for needed improvements related to idling and seat belt usage, and consequent emphasis placed in the training presentation on these two metrics, the most noticeable advantages seen by training group drivers in the initialization followed accordingly. With regards to fuel consumption goals, the training group experienced a 45.9% decrease in idling time per mile driven when compared to the control group. With regards to seat belt usage, the training group experienced a 73.7% decrease in seat belt violations. Although these results are relatively short-term, they are a proof-of-concept and a foundation that indicates constructive changes in driver behavior, which both this specific agency's fleet can propagate and that other agencies can adopt as a model for improvements in safety and eco-driving. And with regards towards future program alterations, more emphasis should be placed on how to limit hard acceleration and cornering while operating the fleet vehicle.

One of the main challenges we experienced in the deployment of the pilot program was mapping data points directly to individual participants. Instead, due to personal security concerns, each data point was linked to a vehicle and our data analysis makes a basic assumption that individual drivers in the training and control groups were the sole vehicle operators for the vehicles selected for study. Although it impacts driver confidentiality, it may be helpful to assign each vehicle a random identifier while maintaining one driver to a singular vehicle. This would help minimize variability and help better model current driving behaviors across a fleet. Additionally, another main challenge in the analysis was building interpretations from the data given the limited number of participants involved.

COVID impacted the team's ability to interact with the drivers in-person, as well as our understanding of current shop procedures and any necessary information the drivers may have given their specific job functions. Furthermore, it may be helpful to identify what routes each driver interacts with the most frequently and receive input as to what training they may like to receive when driving. Although the training was conducted in a virtual environment with somewhat limited interaction, mid-training feedback recorded during the session indicated that drivers were captivated and interactive with the trainer. In the future, it would be useful to conduct training sessions in-person to increase opportunities for driver engagement, minimize possible distractions and provide opportunities for drivers to interact with their vehicles mid-training.

Additional work could be completed to further identify the effectiveness of feedback-based technologies in fleet vehicles as a supplement to the training program. Additionally, because there were only four weeks between implementing the training program and collecting training results, it may be worthwhile to compare the results of training to a long period of time such as 3-6 months after training along with the 4-week window used after training. This would help determine the effectiveness of the training and determine whether or not the training program has both long- and short-term impacts on driver performance.

VI. ACKNOWLEDGMENTS

We thank the FM staff and leadership for supporting this collaborative effort to propel sustainability and safety for drivers. We would also like to thank the drivers involved for their participation in our pilot program.

REFERENCES

[1] Huang, Y., Ng, E. C. Y., Zhou, J., Surawski, N., Chan, E., & Hong, G. (2018). Eco-driving technology for sustainable road transport: A review. Renewable and Sustainable Energy Reviews. 93. 10.1016/j.rser.2018.05.030.

[2] Rakotonirainy, A., Haworth, N., Saint-Pierre, G., & Delhomme, P. (2011). *Research Issues in Eco-Driving* (Queens University of Technology Graduate Dissertation). French Institute in Science and Technology of Transport.

[3] US EPA, O. (2018). Sources of Greenhouse Gas Emissions [Overviews and Factsheets]. US EPA. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions.

[4] Kanaroglou, P. & Buliung, R. (2008). Estimating the contribution of commercial vehicle movement to mobile emissions in urban areas. Transportation Research Part E: Logistics and Transportation Review, 44(2), 260-267. doi.10.1016/j.tre.2007.07.005.

[5] Barkenbus, J. (2010). Eco-Driving: An overlooked climate change initiative. *Energy Policy*, *38*(2), 762–769. https://doi.org/10.1016/j.enpol.2009.10.021

[6] Horrey, W. J., Lesch, M. F., Dainoff, M. J., Robertson, M. M., & Noy, Y. I. (2012). On-Board Safety Monitoring Systems for Driving: Review, Knowledge Gaps, and Framework. Journal of Safety Research, 43(1), 49-58. doi:10.1016/j.jsr.2011.11.004.

[7] Toledo, T., Musicant, O., & Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. Transportation Research Part C: Emerging Technologies, 16(3), 320-331. doi:10.1016/j.trc.2008.01.001.

[8] National Traffic Safety Institute. (2020) SAFER Driver Fleet and Distracted Driver Avoidance Online Course.