NURail Project ID: NURail2016-MTU-R14

Evaluation of Driver Behavior at Railroad-Highway Grade Crossings Using Naturalistic Driving Study Data

By

Pasi Lautala, Ph.D., P.E.
Associate Professor
Director, Rail Transportation Program
Department of Civil and Environmental Engineering
Michigan Technological University
Email: ptlautal@mtu.edu

Myounghoon Jeon, Ph.D.
Associate Professor
Department of Cognitive and Learning Sciences,
Department of Computer Science,
Michigan Technological University,
1400 Townsend Drive, Houghton, MI 49931
Phone: (906) 487-3273
mjeon@mtu.edu

Grant Number: DTRT13-G-UTC52 (Grant 2)
DISCLAIMER

Funding for this research was provided by the NURail Center, University of Illinois at Urbana - Champaign under Grant No. DTRT13-G-UTC52 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research & Technology (OST-R), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation’s University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.
Title
Driver Behavior at Highway-Rail Grade Crossings Using NDS and Driving Simulators

Introduction
What do drivers do in real life as they approach Highway Rail Grade Crossings (HRGCs)?
Highway-Railroad Grade Crossings (HRGCs) are locations where a highway (road or street), including its associated pathways and sidewalks, cross one or more railroad tracks at grade. HRGCs may also be called railroad crossings (RC) or level crossings (LC). HRGCs may be public or private. Private HRGCs are not maintained by public highway authorities and are not intended to be used by the public. According to the U.S. Federal Railroad Administration (FRA), there were total of 211,631 HRGCs operating in the United States in 2015 and more than 60 percent of them were considered public.

Together with trespassing incidents, HRGC accidents (also called collisions or crashes) between roadway vehicles and trains, are the greatest source of injuries and fatalities related to rail transportation in North America. A motorist is 40 times more likely to be killed in a vehicle-train accident than in any other type of highway collision. To illustrate the seriousness of the problem, there were 18,289 collisions between 2008-2017, resulting in 2,250 fatalities and over 8,000 injuries. Because of numerous safety efforts, the total number of HRGC accidents has significantly decreased over the last decades. However, since 2009 the number of HRGC accidents has increased slightly, most likely due to the increased rail and road traffic volumes (Figure 1).
Drivers’ behavior and their reaction to the surrounding conditions and traffic control devices (TCDs) at HRGCs are key elements in both cause and prevention of accidents. FRA’s 2016 report on HRGC accidents states that 94 percent of train-vehicle collisions can be attributed to driver behavior or poor judgment, implying that risky behavior (or lack of defensive driving) by drivers is likely to increase the possibility of an accident at HRGCs. Previous studies on HRGC accidents have also indicated several other factors that increase the accident risk at HRGCs. These factors include rail and highway traffic volumes, train speeds, number of tracks and highway lanes, HRGC angle, TCD type, driver demographics and time of day for the traversal.

The long-standing challenge to lower the number of casualties and accidents at HRGCs warrants a consideration of any new methods and technologies to help in the quest toward zero accidents. We used two potential approaches, naturalistic driving study data and driving simulators. Naturalistic driving studies use instrumented vehicles of everyday drivers to quantitatively evaluate the behavior of those drivers.

Driving simulators are used in a variety of research domains to offer insight into driver behavior. They allow research in controlled environment and hence provide complimentary technology for naturalistic studies. However, it is important to establish the validity of simulator data as a surrogate measure of real-world behavior in a specific given context before extrapolating the results to inform public policy, or the design of new technology. This study uses the correlation
between naturalistic driving study results and simulated data as an example of such validation process.

**Approach and Methodology**

Previous research on HRGC safety has often concentrated on accident reports to predict situations when HRGC accidents are more likely to happen, or used the traffic volumes and infrastructure conditions as an indicator of the risk level at HRGCs. In other words, many past studies on HRGC safety, especially those looking into the role of human behavior, have concentrated on after the fact analysis of accident events. Some other methods, such as external video recordings and roadside or in-vehicle observations have also been used, but those efforts have often provided only partial data of the driving event (internal or external) and tend to have limited sample sizes for developing large-scale trends.

A few past studies have evaluated naturalistic driving data. These naturalistic driving studies have examined motorist behavior by installing video cameras and sensors in automobiles and analyzing the drivers’ actions. For example, the FRA conducted an evaluation of driver behavior at HRGCs in a 2010 study involving light vehicle drivers. The data collected for each grade crossing included information about drivers’ activities, driver and vehicle performance, driving environment, and vehicle location at or on approach to highway-rail grade crossings.

We continue the HRGC safety research that uses direct and detailed observation of the drivers. The overall objective of this two-phase project is to investigate driver behavior at HRGCs using two distinct, but complimentary techniques. Phase 1 of this project takes advantage of the extensive Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) database, collected over several years. The NDS study approach allows for systematic analysis of in-vehicle video and other sensors for direct observation of drivers during typical driving activities at HRGCs. We used the data together with an evaluation methodology developed at Michigan Tech as part of the project in an attempt to quantify the level of defensive driving behavior during HRGC traversals. More specifically, the analysis concentrated on

- driver response to different traffic control devices (TCDs) in place at HRGCs,
- comparison of driver behavior at HRGCs with and without accidents between 2000 and 2010, and
- exploration of the use of NDS data for trending analysis

Phase 2 took advantage of the understanding developed in the first phase to create simulated scenarios that resemble environments found in the NDS data. Driver behavior data from these simulated scenarios was collected and compared with the NDS data sets. More specifically, the objectives were to:

- select two HRGCs from the NDS dataset and recreate them in a simulated setting.
- recruit student drivers to participate in a driving simulation study where drivers are exposed to different HRGCs.
- compare and contrast driver behavior between the NDS and simulator data sets, using the previously developed driver behavior score.

**Findings Phase 1**

We used a three-point driver behavior score methodology developed at Michigan Tech and over 9,000 records of NDS data from the SHRP2 program to quantitatively evaluate driver behavior at HRGCs. Both the use of NDS data for the analysis and scoring methodology were novel and based on drivers visual scanning and speed adjustment while approaching HRGCs. A higher behavior score was considered an indication of a more defensive driving behavior.

**TCD Analysis**

The results from the TCD analysis revealed no statistically significant difference in driver behavior at HRGCs with various types of TCDs, except for HRGCs with passive warnings that included stop signs. In general, the total driver behavior scores in HRGCs with passive warning devices aligned fairly closely with those equipped with active warning devices, even if the visual scanning and speed adjustment components of the score were investigated separately. The average scanning scores of 1.13 suggests that drivers are scanning in both directions during approximately 60% of the traversals, while the speed score of approximately 0.284 suggests that drivers are preparing to stop at 30% of the HRGCs. A closer look at the actual scores revealed that in approximately 34% of the traversals drivers look both ways, 48% look one way, and 18% fail to look at all. These values were consistent across all types of TCDs, except HRGCs with stop signs. Scores for HRGCs with active TCDs showed speed reduction approximately 24% of the time, while the corresponding percentage at passive HRGCs with crossbucks warning only was lower (20%). The results were slightly better for yield controlled HRGCs, with 32% of drivers preparing to stop, but they were based on only seven HRGCs. Overall, we believe that the equal or even lower scores of visual scanning/speed adjustment at passive HRGCs is concerning. While locations with active TCDs provide visual warning of approaching trains to drivers, their only defense at passive HRGCs is through individual recognition of potential danger. The majority of drivers that do not prepare for potential train arrival place themselves at risk during every traversal, providing one explanation for higher accident rates witnessed over time at passive HRGCs.

As mentioned earlier, the only statistically significant exception were higher mean scores at passive HRGSs equipped with stop signs. The overall difference was mainly caused by better performance in the speed reduction category, as the percentage of drivers preparing to stop increased to above 70% when a stop sign was present. This percentage might be even higher, as in some traversals a low entry speed, combined with failure to come to a complete stop, led to an insufficient speed reduction to score a point. We interpret this result as confirmation that drivers express a higher level of defensive driving at HRGCs with stop signs and as such, the requirements
in the MUTCD to increase their usage [19] seem warranted. On the other hand, it remains to be seen how permanent the improvement in driver behavior is, once drivers start to encounter them more frequently.

**Accident Analysis**

Fifteen HRGCs in our NDS database had witnessed accidents between 2000 and 2010, reducing the overall sample size available for the analysis. We found that mean driver behavior scores at accident locations were similar to corresponding values in the full sample. The only notable difference was the higher mean score at the one (the only) passive HRGC with yield sign, but this also highlights the shortcoming of the analysis, i.e. the small number of HRGCs in the accident dataset. Overall, the small sample size makes the results questionable and no conclusions could be made to differentiate driver behavior between HRGCs with and without accidents.

**Trending Analysis**

As noted before, the initial trending analyses were based on looking at parameters individually and their impact on the behavior score. The analyses were broken down to two categories; Category 1 that concentrated on “crossing specific” parameters, and Category 2 that concentrated on “traversal specific” parameters. For Category 1, the analyses were exploratory in nature and not tested for statistical significance.

**Category 1 – Crossing Specific Parameters**

The trending analysis on crossing specific parameters showed several linear trends across the various values. While the statistical significance was not tested, the results can be considered at least as valuable observations that include the following:

- The preliminary trend analysis revealed a moderate decrease in both the scanning and speed scores as AADT increases. One explanation may relate to the fact that as AADT increases the use of active TCDs can also be expected to increase. The added reliance on the active TCDs is likely to reduce the mean driver behavior scores. In addition, with higher AADT drivers are more likely to follow the traffic flow when traversing HRGC.

- Scanning behavior appears to increase slightly as the number of trains per day (TPD) increases. Interestingly, the scan scores for HRGCs with no reported thru trains are very similar to the rest of the HRGCs in our database. Speed scores do appear to increase with train traffic, suggesting that drivers are more likely to reduce their speed with more frequent train traffic.

- There seems to be a clear trend for both lower scanning and speed behavior scores as highway speed limit increases. Just like with AADT, we believe this decline to be related to increased presence of active TCDs on roadways with higher speed limits. However, the results could also be interpreted to suggest that at higher speeds driver attention is more focused on the features within the roadway, leaving less attention for potential trains.
The scanning behavior does not seem to be affected by the train speed at HRGCs. However, the speed score shows increase as train speed increases, suggesting that drivers are more aware of the potential danger from higher speed trains.

Category 2 – Traversal Specific Parameters

While the Category 2 analysis did not reveal any major trends between the scenarios, we found statistically significant differences between several condition pairs tested.

- We found that under different weather conditions, drivers received the highest behavior scores in snow and lowest in rainy conditions (fog was even lower, but the sample size was too small for statistical analysis). Since low driver behavior score was indicative of less defensive driving, the results support the findings from previous studies for increased risk in rainy conditions. However, they do not support the previous finding for increased accident risks under snowy conditions.

- The results were more consistent on differences in driver behavior based on time of day. We found that all drivers from all gender/age groups received significantly lower behavior scores during the night compared with day. This outcome supports previous studies that have revealed poor visibility conditions to have negative impact on driver behavior, thus increasing accident risk at HRGCs. Based on our study results, we suspect that poor visibility at night time lead to higher concentration on the road ahead and as such to lower driver behavior score.

- We did not find any significant difference in average behavior scores of male and female drivers. Based on driver demographic analysis, the only statistically significant difference was between middle-aged female drivers (35-54 years old) when compared to male drivers in the same age category. This differs from earlier studies that concentrated on accidents at HRGCs and found younger male drivers to be at higher risk.

Phase 2 – Simulator Study

We used two existing HRGCs included in the NDS dataset to develop simulated scenarios in the National Advanced Driving Simulator (NADS). We then recruited student drivers to investigate driver behavior at HRGCs in the simulated settings and to compare and contrast driver behavior between the NDS and simulated data sets. The evaluations used the same 3-point quantitative method as the NDS analysis. The following section briefly discusses the key findings from Phase 2 of the study.

- Post-experiment survey revealed that the simulator was considered fairly realistic and participants maintained their driving style. However, the higher operating speeds by Study 1 participants (and the post-experiment survey results) suggest that the speed perception in the simulator differs from that in the natural environment.

- Comparisons between SIM data from Study 1 and Study 2 highlight the effects of train presence and verbal speed instructions, as those were the only differences between the two studies. Participants drove around 80 mph in the simulator without the explicit verbal
instructions to follow the 45mph speed limit. However, we found no statistical difference for mean behaviors scores between stimulation studies 1 and 2, which suggests participants approached HRGC with the same amount of caution regardless of average vehicle speed and the inclusion/exclusion of train present events.

Our main findings from comparing and contrasting the results of NDS and SIM data sets are as follows:

- Both data sets suggest that drivers’ behavior at the crossings have room for improvement from visual scanning and speed adjustment perspective. Mean driver behavior scores are low for both NDS and SIM datasets, 0.8 and 0.6 out of possible 3 total points, respectively. Especially, speed reduction/adjustment behavior is poor in both NDS and SIM datasets. Very few crossing events (26/540, or 4.8% of total) in either data set (NDS and SIM) were awarded a point for speed reduction using current criteria and very few drivers (4/540, or 0.7% of total) received the full 3 points for the driver behavior score.

- A closer look reveals the difference between the two data sets. Statistical tests (ANOVA and t-test) suggested statistically significant difference in the driver behavior scores between NDS and SIM data sets, but not between the different TCDs (Scenario A vs. C).

- The significant main effect for data source suggests participants in the simulation approached HRGCs less cautiously than participants in the NDS dataset. The significant main effect for TCD type suggests drivers approached HRGC with lights/gates more cautiously than HRGC with flashing lights only, but we only observed this trend in the NDS dataset in post-hoc subgroup analyses. Similarly, behavior scores were higher for scenario A (gates/lights) in the NDS data set, but the same trend was not observed in the SIM dataset.

- The post-experiment survey revealed participants’ mixed knowledge and the gap between their knowledge and behavior. Some participants drove in the simulator more cautiously than in their actual driving, while other participants drove less cautiously. Also, half of them reported they would approach active warnings in the off position with caution, but the other half reported they would not slow down or scan for trains because warning is off. All participants understood that it was their responsibility to search for and yield to oncoming trains for passive RR warnings, but (consistent with other driving behavior research) their subjective responses do not correspond with their actual behavior patterns.

**Study Limitations**

Based on our knowledge, this study was the first attempt to use the SHRP2 NDS data for the driver behavior analysis at the HRGCs. We developed a novel “driver behavior score” methodology for the analysis and for the comparisons with simulator experiments conducted in Phase 2. We want to highlight the following limitations and challenges identified in our data and approach during the study.

- **Completeness and accuracy of NDS data**: Due to the difficulties with keeping the sensor arrays in 3,500 vehicles running all the time over an extended period of time, some of our NDS data records had missing or incomplete data. For example, some data records missed
GPS data and numerous records didn’t have data for gas or brake pedal depressions. We have investigated methods for circumventing the missing data, and have developed some techniques to bring some of the excluded records back into our analysis.

- **Use of a 3-point score as the sole qualifier for the analysis:** For simplicity, our methodology relied on developing a single driver behavior score, based on two activities (head rotation and speed reduction) as a quantitative indicator for all behavior during the HRGC traversal. However, condensing a whole chain of events into a single score limits the possibility to investigate the impact of specific factors on driver behavior. It also combines large number of HRGCs with varying characteristic and excludes certain types of HRGCs from the analysis (such as HRGCs near highway intersections). An alternative way to use NDS data in HRGC safety analysis is concentrating on a single behavior, such as the location of speed reduction. Such targeted analysis might provide more granular data for parametric safety analysis.

- **Limited sample size:** While over 9,000 samples is significant amount of data for the analysis, its division into numerous subcategories reduces the sample size per category. Especially the low number of HRGCs with passive TCDs limits the comparative analysis. This is also true on the simulator side. Two HRGCs and 40+ candidates with limited background diversity are at a best “a good start” for statistical analysis.

- **Simulator perception:** Research has shown that the level of “realism” or the fidelity of the driving simulator scenario do not necessarily influence the study outcomes. However, we acknowledge that the perception about the actual risk may be different between the two and can affect the study outcomes.

**Conclusions and Recommendations**

The Phase 1 of this study we used over 9,000 individual traversals obtained from the SHRP 2 naturalistic driving study (NDS) data to evaluate driver behavior at HRGCs. We developed a three-point behavior score based on visual scanning for trains and vehicle speed reduction for the quantitative evaluation. Mean scores were used to generate comparisons and trends, based on selected parameters and conditions. The main comparisons included HRGCs divided by different TCDs and with or without an accident history. In Phase 2, we implemented selected HRGCs from NDS data set in a driver simulator. We used the same evaluation methodology to compare and contrast the simulator results with NDS ones.

While the NDS data analysis resulted in numerous interesting observations, they showed little statistical difference in driving behavior between any of the TCDs analyzed. The only exception was the significantly higher mean scores at passive HRGCs equipped with stop signs. The other consistent finding was the higher mean scores for traversals that took place during the day versus
night time. We found no significant evidence of systematic driver behavior differences in most other categories tested, such as behavior between genders or between age groups.

It was evident from the NDS data (and driver behavior score) that most drivers were not scanning for trains, nor were they preparing to stop, even at crossings with passive warning devices where drivers must rely on their own observations for a safe passage. For example, in the driver simulator scenarios, less than 10 percent of all participants (NDS and simulator data combined) received a point for a proper speed adjustment. With this result, we plan to conduct further research on how to influence drivers to behave appropriately using additional effective but cost-efficient methods (e.g., in-vehicle alerts).

In Phase 2, we used driver simulator experiment and related comparisons to investigate similarities and differences in the driver behavior scores between simulated and natural settings. In general, similar trends could be observed in each data set, but there were more significant differences in scores between scenarios in NDS than in simulated setting, a finding that may be attributed to the relative perception of “safeness” of the simulated environment.

We remain convinced that better understanding driver behavior at HRGCs would be of value when predicting situations when drivers are less cautious and could be at risk of accidents. Despite the limitations and shortcomings of the current effort, we feel confident that NDS data can be beneficial in improving that understanding. We also found no reason to invalidate the use of driver simulators in the analysis. Instead, we believe that they allow researchers to quickly and efficiently analyze new methods to address driver behavior and potential new safety improvements at HRGCs.

The methodologies and data processes developed in this study to evaluate parameters, such as weather, demographics, time of day, etc. are by no means perfect, but they do offer an opportunity to try to quantify the performance of everyday drivers on a large scale. The limitations identified in the report should be addressed and our methodologies modified to improve the accuracy and credibility of the analysis. Instead of relying on a single three-point scale to evaluate the complete behavior, the next steps in the research may concentrate on improving the understanding of individual parameters, such as the exact locations where drivers adjust their speed when approaching HRGC. Research may also consider the use of multivariate analysis techniques to investigate which environmental variables, or groups of variables, have the most significant effects on HRGC behavior. We also believe that value can be found in harnessing machine learning and artificial intelligence applications for predicting driver behavior and their effects on perceived risks.
Publications


A. Dean, D. Nelson, P. Lautala, and M. Jeon. “*Development and Validation of Post-Processing Methods for the SHRP2 MASK Head Pose Data*”. In Transportation Research Circular E-C229: 10th SHRP2 Safety Data Symposium, Washington, D.C., October 6, 2017. ISSN 0097-8515

P. Lautala, D. Nelson, M. Jeon and M. Muhire. “*Using NDS Data to Evaluate Driver Behavior at Highway–Rail Grade Crossings*”. In Transportation Research Circular E-C229: 10th SHRP2 Safety Data Symposium, Washington, D.C., October 6, 2017. ISSN 0097-8515


Primary Contact

**Principal Investigator**

Pasi Lautala, Ph.D., P.E.*

Associate Professor

Director, Rail Transportation Program,

Department of Civil and Environmental Engineering,

Michigan Technological University,

1400 Townsend Drive, Houghton, MI 49931

phone: (906) 487-3547

ptlautal@mtu.edu

Other Faculty and Students Involved

**Myounghoon Jeon, Ph.D.**

Associate Professor

Department of Cognitive and Learning Sciences,

Department of Computer Science,

Michigan Technological University,

1400 Townsend Drive, Houghton, MI 49931

Phone: (906) 487-3273

mjeon@mtu.edu
**David Nelson, P.E.**  
Senior Research Engineer  
Michigan Tech Transportation Institute  
Michigan Technological University,  
1400 Townsend Drive, Houghton, MI 49931  
phone: (906) 487-1734  
dannelso@mtu.edu

**Aaron Dean**  
Undergraduate Student Research Assistant  
Mechanical Engineering  
Michigan Technological University,  
1400 Townsend Drive, Houghton, MI 49931  
phone: (906) 235-8103  
ajdean@mtu.edu

**Modeste Muhire**  
Graduate Research Assistant, Rail Transportation Program,  
Civil and Environmental Engineering Department,  
Michigan Technological University,  
1400 Townsend Drive, Houghton, MI 49931  
Phone: (951) 534-7870  
mmuhire@mtu.edu

**Alawudin Salim**  
Fulbright Scholar  
Civil and Environmental Engineering Department,  
Michigan Technological University,  
1400 Townsend Drive, Houghton, MI 49931  
Phone: (315) 949-8361  
asalim@mtu.edu

**NURail Center**  
217-244-4999  
nurail@illinois.edu  
http://www.nurailcenter.org/
NOTICE
This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof. Any opinions, findings and conclusions, or recommendations expressed in this material do not necessarily reflect the views or policies of the United States Government, nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government. The United States Government assumes no liability for the content or use of the material contained in this document.

NOTICE
The United States Government does not endorse products or manufacturers. Trade or manufacturers’ names appear herein solely because they are considered essential to the objective of this report.
# Driver Behavior at Highway-Rail Grade Crossings Using NDS and Driving Simulators

**Authors:** Pasi Lautala, Myounghoon Jeon, David Nelson, Steven Landry, Aaron Dean

**Performing Organization:**
Michigan Technological University
1400 Townsend Drive
Houghton, MI 49931-1200

**Sponsoring/Monitoring Agency:**
U.S. Department of Transportation
Federal Railroad Administration
Office of Railroad Policy and Development
Office of Research and Development
Washington, DC 20590

**Abstract:**
Michigan Technological University (Michigan Tech) used the Second Strategic Highway Research Program (SHRP2) naturalistic driving study (NDS) data and driver simulator to perform a quantitative evaluation of driver behavior at highway rail grade crossings (HRGCs). We developed a three-point scale to generate driver behavior score for over 9,000 NDS traversals and used the mean scores to perform statistical comparisons of driver behavior at HRGCs with different traffic control devices (TCDs), with/without accident history, and with various environmental conditions. We also simulated two HRGCs in a driver simulator and compared the driver behavior scores between naturalistic and simulated environment.

The investigation revealed that most drivers do not visually scan for trains and do not prepare to stop, regardless the type of warning device present at the crossing, or the environmental conditions that prevail at the time of traversal. The results were fairly consistent in both NDS and simulated approaches. The NDS data analysis showed very little statistical difference in driving behavior between any of the TCDs analyzed. The only exceptions were the significantly higher mean scores at passive HRGCs equipped with stop signs and the higher mean scores for traversals that took place during the day versus night time.

**Subject Terms:**
Grade crossing, traffic control device, average annual daily traffic, Federal Highway Administration, Federal Railroad Administration, vehicle miles traveled, freight train miles, fatality analysis, reporting system, Daylight Saving Time, data normalization, crossbuck, traffic control device

**Number of Pages:**
xxx

## METRIC/ENGLISH CONVERSION FACTORS

### ENGLISH TO METRIC

<table>
<thead>
<tr>
<th>LENGTH (APPROXIMATE)</th>
<th>METRIC TO ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 inch (in) = 2.5 centimeters (cm)</td>
<td>1 millimeter (mm) = 0.04 inch (in)</td>
</tr>
<tr>
<td>1 foot (ft) = 30 centimeters (cm)</td>
<td>1 centimeter (cm) = 0.4 inch (in)</td>
</tr>
<tr>
<td>1 yard (yd) = 0.9 meter (m)</td>
<td>1 meter (m) = 3.3 feet (ft)</td>
</tr>
<tr>
<td>1 mile (mi) = 1.6 kilometers (km)</td>
<td>1 meter (m) = 1.1 yards (yd)</td>
</tr>
</tbody>
</table>

### AREA (APPROXIMATE)

<table>
<thead>
<tr>
<th>ENGLISH TO METRIC</th>
<th>METRIC TO ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 square inch (sq in, in^2) = 6.5 cm^2</td>
<td>1 square centimeter (cm^2) = 0.16 sq in</td>
</tr>
<tr>
<td>1 square foot (sq ft, ft^2) = 0.09 m^2</td>
<td>1 square meter (m^2) = 1.2 sq ft</td>
</tr>
<tr>
<td>1 square mile (sq mi, mi^2) = 2.6 km^2</td>
<td>1 square kilometer (km^2) = 0.4 sq mi</td>
</tr>
<tr>
<td>1 acre = 0.4 hectare (he) = 4,000 m^2</td>
<td>10,000 square meters (m^2) = 1 hectare (ha) = 2.5 acres</td>
</tr>
</tbody>
</table>

### MASS - WEIGHT (APPROXIMATE)

<table>
<thead>
<tr>
<th>ENGLISH TO METRIC</th>
<th>METRIC TO ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ounce (oz) = 28 grams (gm)</td>
<td>1 gram (gm) = 0.036 ounce (oz)</td>
</tr>
<tr>
<td>1 pound (lb) = 0.45 kilogram (kg)</td>
<td>1 kilogram (kg) = 2.2 pounds (lb)</td>
</tr>
<tr>
<td>1 short ton = 2,000 pounds (lb) = 0.9 tonne (t)</td>
<td>1 tonne (t) = 1.000 kilograms (kg) = 1.1 short tons</td>
</tr>
</tbody>
</table>

### VOLUME (APPROXIMATE)

<table>
<thead>
<tr>
<th>ENGLISH TO METRIC</th>
<th>METRIC TO ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 teaspoon (tsp) = 5 milliliters (ml)</td>
<td>1 milliliter (ml) = 0.03 fluid ounce (fl oz)</td>
</tr>
<tr>
<td>1 tablespoon (tbsp) = 15 milliliters (ml)</td>
<td>1 liter (l) = 2.1 pints (pt)</td>
</tr>
<tr>
<td>1 fluid ounce (fl oz) = 30 milliliters (ml)</td>
<td>1 liter (l) = 1.06 quarts (qt)</td>
</tr>
<tr>
<td>1 cup (c) = 0.24 liter (l)</td>
<td>1 liter (l) = 0.26 gallon (gal)</td>
</tr>
<tr>
<td>1 pint (pt) = 0.47 liter (l)</td>
<td></td>
</tr>
<tr>
<td>1 quart (qt) = 0.96 liter (l)</td>
<td></td>
</tr>
<tr>
<td>1 gallon (gal) = 3.8 liters (l)</td>
<td></td>
</tr>
<tr>
<td>1 cubic foot (cu ft, ft^3) = 0.03 cubic meter (m^3)</td>
<td>1 cubic meter (m^3) = 36 cubic feet (cu ft, ft^3)</td>
</tr>
<tr>
<td>1 cubic yard (cu yd, yd^3) = 0.76 cubic meter (m^3)</td>
<td>1 cubic meter (m^3) = 1.3 cubic yards (cu yd, yd^3)</td>
</tr>
</tbody>
</table>

### TEMPERATURE (EXACT)

<table>
<thead>
<tr>
<th>ENGLISH TO METRIC</th>
<th>METRIC TO ENGLISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>[(x-32)(\times\frac{5}{9})] F = (\frac{9}{5}) y + 32</td>
<td>[(9/5) y + 32] C = x (\times) F</td>
</tr>
</tbody>
</table>

### QUICK INCH - CENTIMETER LENGTH CONVERSION

<table>
<thead>
<tr>
<th>Inches</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centimeters</td>
<td>0</td>
<td>2.5</td>
<td>5</td>
<td>7.5</td>
<td>10</td>
<td>12.5</td>
</tr>
</tbody>
</table>

### QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSION

<table>
<thead>
<tr>
<th>°F</th>
<th>-40°</th>
<th>-22°</th>
<th>-4°</th>
<th>14°</th>
<th>32°</th>
<th>50°</th>
<th>68°</th>
<th>86°</th>
<th>104°</th>
<th>122°</th>
<th>140°</th>
<th>158°</th>
<th>176°</th>
<th>194°</th>
<th>212°</th>
</tr>
</thead>
<tbody>
<tr>
<td>°C</td>
<td>-40°</td>
<td>-30°</td>
<td>-20°</td>
<td>-10°</td>
<td>0°</td>
<td>10°</td>
<td>20°</td>
<td>30°</td>
<td>40°</td>
<td>50°</td>
<td>60°</td>
<td>70°</td>
<td>80°</td>
<td>90°</td>
<td>100°</td>
</tr>
</tbody>
</table>

For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98
Acknowledgements

Research team would like to acknowledge the following individuals and organizations for their support and contributions to the project:

- Starr Kidda and Debra Chappell, Federal Railroad Administration for technical support
- Miguel Perez, Edie Sears, and the rest of the team at Virginia Tech Transportation Institute for support with SHRP2 NDS data
- The National Advanced Driving Simulator (NADS) team, Iowa State University
- Michigan Tech Transportation Institute
- National University Rail Center (NURail), USDOT Tier-1 University Transportation Center
- Student researchers at Michigan Tech (in addition to contributing authors); Kyle Dick, Modeste Muhire, Darian Reed, Alawudin Salim and Yugang Wang
# Contents

Acknowledgements.................................................................................................................. iii
Illustrations ................................................................................................................................. v
Tables ........................................................................................................................................ vii
Executive Summary .................................................................................................................... 8
1. Introduction ............................................................................................................................. 9
   1.1 Background ....................................................................................................................... 9
   1.2 Study Objectives and Scope .......................................................................................... 11
   1.3 Organization of the Report ............................................................................................ 12
2. Phase 1 – Data Sources and Methodology ............................................................................. 13
   2.1 Data Sources .................................................................................................................. 14
   2.2 Data Acquisition and Processing ................................................................................... 16
   2.3 Statistical Analysis .......................................................................................................... 20
3. Results and Data Analysis ....................................................................................................... 22
   3.1 Summary of Behavior Scores ....................................................................................... 22
   3.2 Summary of Statistical Test Results for TCDs and at Accident Locations .................... 23
   3.3 Trending Parameter Analysis ......................................................................................... 25
4. Discussion and Limitations of Phase 1 ................................................................................... 35
   4.1 TCD Analysis ................................................................................................................. vi
   4.2 Accident Analysis ............................................................................................................. vii
   4.3 Trending Analysis .............................................................................................................. vii
5. Phase 2 – Data Sources and Methodology for Driver Simulator Study ................................. 38
   5.2 Results and Data Analysis ............................................................................................ 41
   5.3 Discussion of Phase 2 ..................................................................................................... 48
6. Study Limitations ..................................................................................................................... 50
7. Conclusion and Recommendation for Future Work .............................................................. 51
8. References ............................................................................................................................... 53
Appendix A. Percent Behavior Scores ....................................................................................... 55
Appendix B. Driving Simulator Description ............................................................................... 58
Illustrations

Figure 1. Number of Annual HRGC Injuries and Fatalities, 2008-2017........................................ iv
Figure 2 - Phase 1 Data Sources and Research Process................................................................. 13
Figure 3 – HRGC Selection Process............................................................................................... 16
Figure 4 - Rotation and Pitch Thresholds Used in Head Tracking................................................. 18
Figure 5 – Driver Behavior Score Analysis Zone............................................................................ 18
Figure 6 - Full (3-point) Behavior Score Example .......................................................................... 19
Figure 7 – “Zero” Behavior Score Example .................................................................................... 20
Figure 8 – Mean Driver Behavior Scores by TCD for Accident and Non-Accident HRGCs .......... 22
Figure 9 – Visual Scan and Speed Scores by Annual Average Daily Traffic (Sample Size Per Category on the Top) ........................................................................................................... 26
Figure 10 – Scan, Speed and Total Scores by Trains Per Day (Sample Size Per Category on the Top) .................................................................................................................................................. 27
Figure 11 - Behavior Scores by Posted Highway Speed (Sample Size Per Category on the Top) ................................................................................................................................................. 28
Figure 12 - Behavior Scores by Train Speed (Sample Size Per Category on the Top) ..................... 29
Figure 13 – Effect of Weather on Mean Driver Behavior Score ...................................................... 30
Figure 14 - Drivers’ Mean Behavior Score Based on Time of Day .................................................. 31
Figure 15 – Mean Driver Behavior Scores by Gender ...................................................................... 32
Figure 16 – Mean Driver Behavior Scores by Age and Sex .............................................................. 33
Figure 17 – Track Designs for Scenario A (left) and C (right). Red Lines Indicate Railroads That Intersect with Highways (Grey). ........................................................................................................ 40
Figure 18 – Simulated advance warning (left) and gated crossing (right) in scenario A .......... 40
Figure 19 – Simulated Advance Warning in Scenario C (Left) Compared to Reference NDS Video (Right) ......................................................................................................................................... 40
Figure 20 – Mean Behavior Score by Data Source (NDS vs. SIM) .................................................. 42
Figure 21 – Mean Behavior Score by Scenario (TCD Type, Scenario A vs. C) .................................. 42
Figure 22 – Mean Behavior Score by Data Source For Scenarios A and C ...................................... 44
Figure 23 – Interaction Plot by Data Source For Scenarios A and C ................................................ 44
Figure 24 – Mean Behavior Score Grouped by NDS and SIM (Study 1 and Study 2) Data Sets 45
Figure 25 – Mean Behavior Score Grouped by NDS, Study 1, and Study 2 Data Sets............... 45
Figure 26 – Stacked Bar Chart Depicting the Contribution From Behavior Score Sub-Behaviors ........................................................................................................................................ 46
Figure 27 – Pie Graphs of Exposure to Crossings and Trains in Real World Driving ............ 48
Figure 28 - Percent of Traversals with the Given Total Score, by TCD.......................................... 56
Figure 29 - Percentage of Traversals with the Given Scan Score by TCD........................................ 57
Figure 30 - Percentage of Traversals with the Given Speed Score by TCD........................................ 57
Figure 31 – NADS MiniSim Driving Simulator.................................................................................. 58
Tables

Table 1 - Time Series data from NDS Trip Database .................................................. 14
Table 2 - Data collected from NDS video files ............................................................ 14
Table 3 - Number of HRGCs per State in the RID ......................................................... 15
Table 4 - Data collected from the FRA Crossing Inventory Database .......................... 15
Table 5 - Data collected from the FRA Accident/Incident Database .......................... 15
Table 6 - Driver Behavior Score Calculation .............................................................. 17
Table 7 – Mean Driver Behavior scores by TCD ......................................................... 23
Table 8 – Mean Driver Behavior scores for Accident Locations ................................. 23
Table 9 - Welch's T-test for comparison of TCD conditions ....................................... 24
Table 10 - Welch's t-test results for comparing accident to non-accident locations .... 25
Table 11 - Summary of Statistical Test Results Based on Weather and Time of Day .... 34
Table 12 - Summary of Statistical Test Results Based on Gender and Age ................ 34
Table 13 - Description of NDS vs. Simulator Data ...................................................... 38
Table 14 – 2x3 Analysis of Variance ............................................................................. 41
Table 15 - Post-hoc Tukey HSD Analysis .................................................................. 45
Table 16 - Percent of total behavior score ................................................................. 47
Table 17 - Count and percent of traversals with speed reduction during the traversal ... 47
Table 18 - Behavior Score by Percent ......................................................................... 55
Executive Summary

Michigan Technological University (Michigan Tech) used the Second Strategic Highway Research Program (SHRP2) naturalistic driving study (NDS) data base and driver simulator to perform a quantitative evaluation of driver behavior at highway rail grade crossings (HRGCs). The purpose of the study was to quantify the level of defensive driving behavior during HRGC traversals. We developed a three-point driver behavior score and automatic data processing application to generate scores for over 9,000 traversals. We used the mean scores to perform statistical comparisons of driver behavior at HRGCs with different traffic control devices (TCDs) and between HRGCs that have and have not witnessed accidents in the past. We also explored whether any trending could be identified among parameters identified as critical to safety in previous studies. Finally, we simulated two HRGCs in a driver simulator and compared the driver behavior scores between naturalistic and simulated environment.

The investigation of over 9,000 NDS traversals revealed that most drivers do not visually scan for trains and do not prepare to stop, regardless the type of warning device present at the crossing, or the environmental conditions that prevail at the time of traversal. The results were fairly consistent in both NDS and simulated approaches.

The main findings of the study included the following:

- The NDS data analysis showed little statistical difference in driving behavior between any of the TCDs analyzed. The only exceptions were the significantly higher mean scores at passive HRGCs equipped with stop signs and consistently higher mean scores for traversals that took place during the day versus night time.
- We found no significant evidence of systematic driver behavior differences when comparing different demographics, such as behavior between genders or between age groups.
- Trending analysis based on average annual daily traffic (AADT), trains per day and train/highway speed impact on driver behavior score provided interesting observations, but statistical analysis were not conducted due to the exploratory nature of the analysis.
- Simulator testing and NDS data analysis provided similar results, but there were more significant differences between scenarios in the NDS than in the simulated setting. This finding may be attributed to the relative perception of “safeness” of the simulated environment.

Despite the limitations and shortcomings of the current effort, we remain convinced that better understanding driver behavior at HRGCs would be of value when predicting situations when drivers are less cautious and could be at risk of accidents. The availability of SHRP2 NDS data base can be beneficial in improving that understanding. However, instead of relying on a simple three-point scale to evaluate the complete behavior, the direction of future research should concentrate on analyzing specific parameters, or parameter clusters, that are expected to have the greatest impact on the behavior. We also found no reason to invalidate the use of driver simulators as part of the analysis, but we also believe that techniques, such as multivariate analysis, machine learning, and artificial intelligence can be harnessed to assist in such research.
• Introduction

Background
Highway-Railroad Grade Crossings (HRGCs) are locations where a highway (road or street), including its associated pathways and sidewalks, cross one or more railroad tracks at grade. HRGCs may also be called railroad crossings (RC) or level crossings (LC). HRGCs may be public or private. Private HRGCs are not maintained by public highway authorities and are not intended to be used by the public. According to the U.S. Federal Railroad Administration (FRA), there were total of 211,631 HRGCs operating in the United States in 2015 and more than 60 percent of them were considered public [1].

Together with trespassing incidents, HRGC accidents (also called collisions or crashes) between roadway vehicles and trains, are the greatest source of injuries and fatalities related to rail transportation in North America. A motorist is 40 times more likely to be killed in a vehicle-train accident than in any other type of highway collision [2]. To illustrate the seriousness of the problem, there were 18,289 collisions between 2008-2017, resulting in 2,250 fatalities and over 8,000 injuries [3]. Because of numerous safety efforts, the total number of HRGC accidents has significantly decreased over the last decades. However, since 2009 the number of HRGC accidents has increased slightly, most likely due to the increased rail and road traffic volumes (Figure 1)[3].
Drivers’ behavior and their reaction to the surrounding conditions and traffic control devices (TCDs) at HRGCs are key elements in both cause and prevention of accidents. FRA’s 2016 report on HRGC accidents states that 94 percent of train-vehicle collisions can be attributed to driver behavior or poor judgment, implying that risky behavior (or lack of defensive driving) by drivers is likely to increase the possibility of an accident at HRGCs [4]. Previous studies on HRGC accidents have also indicated several other factors that increase the accident risk at HRGCs. These factors include rail and highway traffic volumes, train speeds, number of tracks and highway lanes, HRGC angle, TCD type, driver demographics and time of day for the traversal [1].

The long-standing challenge to lower the number of casualties and accidents at HRGCs warrants a consideration of any new methods and technologies to help in the quest toward zero accidents. We used two potential approaches, naturalistic driving study data and driving simulators. Naturalistic driving studies use instrumented vehicles of everyday drivers to quantitatively evaluate the behavior of those drivers.

Driving simulators are used in a variety of research domains to offer insight into driver behavior. They allow research in controlled environment and hence provide complimentary technology for naturalistic studies. However, it is important to establish the validity of simulator data as a surrogate measure of real-world behavior in a specific given context before extrapolating the results to inform public policy, or the design of new technology. This study uses the correlation between naturalistic driving study results and simulated data as an example of such validation process.
Study Objectives and Scope

Previous research on HRGC safety has often concentrated on accident reports to predict situations when HRGC accidents are more likely to happen, or used the traffic volumes and infrastructure conditions as an indicator of the risk level at HRGCs. In other words, many past studies on HRGC safety, especially those looking into the role of human behavior, have concentrated on after the fact analysis of accident events [5,6]. Some other methods, such as external video recordings and roadside or in-vehicle observations have also been used, but those efforts have often provided only partial data of the driving event (internal or external) and tend to have limited sample sizes for developing large-scale trends.

A few past studies have evaluated naturalistic driving data. These naturalistic driving studies have examined motorist behavior by installing video cameras and sensors in automobiles and analyzing the drivers’ actions. For example, the FRA conducted an evaluation of driver behavior at HRGCs in a 2010 study involving light vehicle drivers. The data collected for each grade crossing included information about drivers’ activities, driver and vehicle performance, driving environment, and vehicle location at or on approach to highway-rail grade crossings. [7].

We continue the HRGC safety research that uses direct and detailed observation of the drivers. The overall objective of this two-phase project is to investigate driver behavior at HRGCs using two distinct, but complimentary techniques. Phase 1 of this project takes advantage of the extensive Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) database, collected over several years [8]. The NDS study approach allows for systematic analysis of in-vehicle video and other sensors for direct observation of drivers during typical driving activities at HRGCs. We used the data together with an evaluation methodology developed at Michigan Tech as part of the project in an attempt to quantify the level of defensive driving behavior during HRGC traversals. More specifically, the analysis concentrated on

- driver response to different traffic control devices (TCDs) in place at HRGCs,
- comparison of driver behavior at HRGCs with and without accidents between 2000 and 2010, and
- exploration of the use of NDS data for trending analysis

Phase 2 took advantage of the understanding developed in the first phase to create simulated scenarios that resemble environments found in the NDS data. Driver behavior data from these simulated scenarios was collected and compared with the NDS data sets. More specifically, the objectives were to:

- select two HRGCs from the NDS dataset and recreate them in a simulated setting.
- recruit student drivers to participate in a driving simulation study where drivers are exposed to different HRGCs.
- compare and contrast driver behavior between the NDS and simulator data sets, using the previously developed driver behavior score.
**Organization of the Report**

This report first summarizes and discusses the research data, activities and outcomes for Phase 1 tasks, followed by a similar summary for Phase 2. Study limitations, conclusions and discussion of potential future research complete the report.
• Phase 1 – Data Sources and Methodology

Figure 2 outlines the data sources and process flow for Phase 1 research. Each process component is explained in more detail in the following sections.

Figure 3 - Phase 1 Data Sources and Research Process
Data Sources

The study used four main data sources; the SHRP 2 Naturalistic Driving Study database (NDS) [8], SHRP 2 Roadway Information Database (RID) [9], the FRA Grade Crossing Inventory Database [10] and FRA Accident/Incident Database [11]. Google Maps, Google Earth and forward video streams of the NDS data were used to verify TCDs at selected HRGCs.

**SHRP 2 NDS and SHRP 2 RID Databases**

The NDS, funded through the Transportation Research Board (TRB) under the National Academies of Science (NAS), captured unsupervised driving performance of participants from six different states in the United States; Florida, Indiana, New York, North Carolina, Pennsylvania and Washington. The study was conducted in 2010-2013 and included more than 5 million trips by approximately 3,500 participants [8]. The data from the NDS trip database used for this study included detailed sensor information on vehicle location, brake and throttle position, vehicle speed and acceleration, and driver demographics (Table 1). It also included front and rear video feeds from the NDS video files, and head rotation and position derived from a face video feed (Table 2). The database is stored in a secure data enclave at the Virginia Tech Transportation Institute (VTTI) and is made accessible for researchers across the US through a Data Use License (DUL).

<table>
<thead>
<tr>
<th>Table 1 - Time Series data from NDS Trip Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Gender</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2 - Data collected from NDS video files</th>
</tr>
</thead>
<tbody>
<tr>
<td>14. Other Features</td>
</tr>
<tr>
<td>15. Crossing Conditions</td>
</tr>
<tr>
<td>17.</td>
</tr>
</tbody>
</table>

The Roadway Information Database (RID) was also developed under the SHRP2 to provide roadway and route information on trips taken by the NDS participants [9]. Michigan Tech used the RID to identify 1,017 public HRGCs traversed by the NDS study participants (Table 3) and selected specific HRGCs from the sample for the analysis in our project.
### Table 3 - Number of HRGCs per State in the RID

<table>
<thead>
<tr>
<th>1. State</th>
<th>2. Number of HRGCs in NDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Florida</td>
<td>4. 295</td>
</tr>
<tr>
<td>5. Indiana</td>
<td>6. 104</td>
</tr>
<tr>
<td>7. New York</td>
<td>8. 181</td>
</tr>
<tr>
<td>11. Pennsylvania</td>
<td>12. 61</td>
</tr>
<tr>
<td><strong>15. Total</strong></td>
<td><strong>16. 1,017</strong></td>
</tr>
</tbody>
</table>

**FRA Grade Crossing Inventory Database and Accident/Incident Database**

Each of the over 200,000 HRGCs in the FRA inventory database is identified by its FRA crossing ID and described by different data fields that provide HRGC information ranging from ownership to field configuration. We used crossing ID as a linking field in a programming algorithm to match the available data in the FRA inventory with the corresponding HRGCs in the NDS database. The absence/presence of active warning devices (lights and/or lights and gates) were used to categorize HRGCs based on their TCDs.

The FRA accident/incident database includes information about reported accidents that have occurred at HRGCs. This source includes several decades’ worth of historical data and was used to identify HRGCs with both an accident history between 2000 and 2010 and traversal data in the NDS. The pertinent data fields used in the study from the two FRA databases are presented in Tables 4 and Table 5.

### Table 4 - Data collected from the FRA Crossing Inventory Database

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>25. Train Traffic Levels</td>
<td>26. Latitude and Longitude</td>
<td>27. Number of Tracks</td>
</tr>
</tbody>
</table>

### Table 5 - Data collected from the FRA Accident/Incident Database
Google Maps and Forward Cameras and Selection of HRGCs for Analysis

We used Google Maps and Google Earth to perform initial verification of TCD status at all 1,017 HRGCs during the NDS study period. We later confirmed the crossing information using the forward-facing video segments of traversals.

Data Acquisition and Processing

We used information from the RID and from the FRA Crossing Inventory database to select diverse group of HRGCs for the analysis in our study. Several key parameters were used in the selection process, including type of TCDs, configuration of nearby intersections and the number of accidents in recent years. Since there was a limited number of passive HRGCs with traversals available in the NDS, all passive HRGCs in all six states were selected for the analysis. We then limited the study area to three states, New York, Indiana, and Florida and selected all 55 HRGCs with flashing lights but no gates for the analysis. For HRGCs with lights and gates, there were over 400 potential candidates, so we used proximity to roadway intersections, HRGC angle, and number of accidents at the HRGC as the main selection criteria. The final list selected for the analysis contained a total of 306 HRGCs, 199 with lights and gates, 55 with flashing lights only, and 54 with passive warning devices (Figure 3).

![Figure 4 - HRGC Selection Process](image)

The requested data set from the NDS database included an average sample of approximately 40 individual traversals per HRGC. This number was calculated to provide statistically valid results, as described in more detail in a paper by Muhire et al, [13].

Once the HRGCs and sample sizes were determined, the final data request from the NDS data archive included data from almost 13,000 individual traversals. We submitted the data request in three batches. This allowed us to begin analysis work earlier and modify the later requests to
achieve a more representative sample. We linked the traversal data to the specific HRGC and stored the compiled data set in an integrated database for further analysis.

**Developing the Behavior Score**

Earlier research at Michigan Tech developed a methodology to quantitatively evaluate driver behavior when approaching a HRGC [14]. Although the evaluation originally used the term “compliance score”, it was later changed to “behavior score”, a term considered more appropriate for the analysis.

The behavior score attempted to translate qualitative driver behavior data into a quantitative “3-point behavior score” suitable for statistical analysis by quantifying two types of actions (Table 6). A higher behavior score was considered an indication of a more defensive driving behavior.

**Table 6 - Driver Behavior Score Calculation**

<table>
<thead>
<tr>
<th>17. Driver Behavior Action</th>
<th>18. Points Awarded</th>
</tr>
</thead>
<tbody>
<tr>
<td>19. Visually scan for train to the right (&gt; 8 degrees)</td>
<td>20. +1</td>
</tr>
<tr>
<td>21. Visually scan for train to the left (&gt; 8 degrees)</td>
<td>22. +1</td>
</tr>
<tr>
<td>23. Identifiable speed adjustment (&gt; 10% reduction)</td>
<td>24. +1</td>
</tr>
<tr>
<td><strong>25. Total Possible Score</strong></td>
<td><strong>26. +3</strong></td>
</tr>
</tbody>
</table>

We used head rotation and position to evaluate visual train scanning behavior. Since a train can typically come from either direction at an HRGC, one point was awarded for scanning in each direction. Scanning scores were based on the headtracking data from the NDS that provides horizontal and vertical head rotation and position data. We used smoothing and interpolation techniques to fill in areas where the original NDS data had gaps [15]. A scan was considered successful if the driver’s head rotation exceeded eight degrees from the baseline after the smoothing (typically the roadway center) in either horizontal direction. To ensure the driver was not looking at things inside the car, the scan was only awarded points, if a pitch, or vertical scan, did not exceed eight degrees from the baseline (Figure 4). More details of the scanning score process can be found in papers by Lautala et. al. [16,17]
Figure 5 - Rotation and Pitch Thresholds Used in Head Tracking

The second component of behavior score was based on an identifiable speed adjustment that indicated driver’s recognition of approaching a HRGC and preparation to stop, as necessary. We tested three different methods to determine the speed adjustment. The early analysis identified instances where the speed profile showed a deceleration of 2 ft/s² within a specific approach zone before the HRGC. We also investigated the pedal movements as an indicator, where score was based on removing the foot from the gas pedal and applying the brake. This proved problematic, as such data was not consistently available from the NDS data.

We finally settled on using a third alternative for the analysis, defined as a minimum 10% speed reduction within the analysis zone. The analysis zone was determined for each traversal based on the stopping distance required at the entry speed (the drivers’ speed as they approached the HRGC), and the reaction time obtained from the Manual for Uniform Traffic Control Devices (MUTCD) (Figure 5) [18]. We first identified the last possible braking point for a driver to safely come to a complete stop before reaching HRGC. The braking point was based on a vehicle deceleration of 11 ft/s² per the MUTCD and also became the end point for the analysis range. We allowed drivers five seconds before the braking point to complete the proper actions (head rotation and speed adjustment), demonstrating their preparation for the coming HRGC traversal. The MUTCD suggests using a reaction time of 2.5 seconds from when the driver can see an obstruction until the driver takes an action in response to the obstruction. We provided drivers with twice that time period to both recognize the potential danger and to react with necessary precautions.

Figure 6 – Driver Behavior Score Analysis Zone

Figure 6 and Figure 7 provide examples demonstrating the calculation of the behavior score. In Figure 6 the top graph shows a decrease in velocity of nearly 10 mph (>20 percent), and the middle graph shows a scan to the right (approximately nine seconds before arrival to the HRGC), and a scan to the left (approximately seven seconds before). The bottom graph shows that the vertical head position stays within the eight degree range during both horizontal scans. In this case, the driver receives a full “three” points for the traversal.
Figure 7 shows little change in speed during the analysis period, and the horizontal head rotation stays inside the eight degree window, suggesting no visual scanning. The pitch would be acceptable, but this traversal would be consistent with a driver looking forward at the road ahead during the entire analysis range. The behavior score for this traversal is “zero”.

Figure 7 - Full (3-point) Behavior Score Example
Statistical Analysis

As described in the project scope, the main objective of the study was to compare driver behavior scores at locations with different TCDs. We also compared behavior between accident and non-accident locations. We used mean driver behavior scores at each HRGC as the primary value for comparisons and conducted a series of paired Welch’s t-tests to identify whether the means were statistically different.

Statistical validity is important when comparing sample scores between different categories of results. In this work we first used statistical techniques and the results from previous small scale studies to establish a target sample size of 60 HRGCs for each type of TCD in the analysis. This sample size was determined sufficient to verify the statistical validity of the results at the 90% confidence level.
We should note that all crossings with the same type of TCD were considered “equal” in this study. This means that other potential differences, such as urban/rural location, number of highway lanes, highway speeds, etc. were not used to categorize the HRCSs in smaller subgroups for the TCD and accident analysis. Although it reduced the homogeneity of the HRGCs in each category, it also ensured as large as possible samples sizes.
• Results and Data Analysis

We reviewed over 12,000 traversals from 306 crossings in the analysis. However, the records from numerous traversals proved unusable due to incomplete data, such as the missing GPS location of the vehicle. After data reduction to remove such records, 9,128 traversals from 286 different HRGCs were determined to have a sufficient level of data to reliably conduct the analysis. The results presented in this section, as well as the statistical test results, are based on this final set of traversals. The results include the mean behavior scores calculated for the TCD groupings, including separate calculations for the accident locations. The results also provide a breakdown between visual scanning and speed adjustment scores and the sample sizes of each category. We also present the results of exploratory trending analysis that use highway and train speeds and highway and train traffic volumes as main parameters, although they did not include statistical analysis. Summary of results from a parallel study that investigated the effects of environmental conditions and demographics on driver behavior are also provided.

Summary of Behavior Scores

Figure 8 shows the mean behavior scores in each TCD category for all crossings and separately for the accident locations. Mean scores between accident and non-accident locations were fairly similar across most categories, except for passive HRGCs with crossbuck and yield signs. However, the accident location data set for this condition included only one HRGC.

Table 7 and Table 8 present the mean driver behavior scores, categorized by the prevailing TCDs at the HRGC during the NDS data collection. The mean values and corresponding standard deviations are provided separately for the visual scan and the speed adjustment and combined in the total score. The number of crossings included in the category is presented in the final column. The results show that the mean speed and visual scan scores were similar within each main category (active vs. passive TCDs), except for the stop conditions.
Table 7 – Mean Driver Behavior scores by TCD

<table>
<thead>
<tr>
<th>27. TCD Type</th>
<th>28. Scan Score</th>
<th>29. Speed Score</th>
<th>30. Total Score</th>
<th>31. # HRGCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall - All HRGCs</td>
<td>33.157</td>
<td>34.0238</td>
<td>35.1395</td>
<td>36.286</td>
</tr>
<tr>
<td>Std Dev</td>
<td>38.0200</td>
<td>39.0231</td>
<td>40.0340</td>
<td>41.0</td>
</tr>
<tr>
<td>Gated HRGC</td>
<td>43.1148</td>
<td>44.0233</td>
<td>45.1381</td>
<td>46.205</td>
</tr>
<tr>
<td>Std Dev</td>
<td>48.0193</td>
<td>49.0229</td>
<td>50.0329</td>
<td>51.0</td>
</tr>
<tr>
<td>Lights, No Gates</td>
<td>53.1153</td>
<td>54.0243</td>
<td>55.1395</td>
<td>56.51</td>
</tr>
<tr>
<td>Std Dev</td>
<td>58.0202</td>
<td>59.0233</td>
<td>60.0358</td>
<td>61.0</td>
</tr>
<tr>
<td>Crossbuck with Yield</td>
<td>63.1189</td>
<td>64.0329</td>
<td>65.1519</td>
<td>66.7</td>
</tr>
<tr>
<td>Std Dev</td>
<td>68.0669</td>
<td>69.0366</td>
<td>70.0818</td>
<td>71.0</td>
</tr>
<tr>
<td>Crossbuck Only</td>
<td>73.1191</td>
<td>74.0239</td>
<td>75.1429</td>
<td>76.23</td>
</tr>
<tr>
<td>Std Dev</td>
<td>78.0176</td>
<td>79.0248</td>
<td>80.0355</td>
<td>81.0</td>
</tr>
<tr>
<td>Crossbuck with Stop</td>
<td>83.1348</td>
<td>84.0745</td>
<td>85.209</td>
<td>86.5</td>
</tr>
<tr>
<td>Std Dev</td>
<td>88.0247</td>
<td>89.0142</td>
<td>90.016911</td>
<td>91.0</td>
</tr>
</tbody>
</table>

Table 8 shows the behavior scores for the HRGCs that had reported accidents between 2000 and 2013. Since the sample size of some categories was small, the value of the statistical analysis was also reduced. HRGCs with only flashing lights and no gates are excluded, as our data set had no accidents in such locations.

Table 8 – Mean Driver Behavior scores for Accident Locations

<table>
<thead>
<tr>
<th>TCD Type</th>
<th>Scan Score</th>
<th>Speed Score</th>
<th>Total Score</th>
<th># HRGCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall - All HRGCs</td>
<td>1.130</td>
<td>0.284</td>
<td>1.403</td>
<td>15</td>
</tr>
<tr>
<td>Std Dev</td>
<td>0.704</td>
<td>0.355</td>
<td>0.859</td>
<td></td>
</tr>
<tr>
<td>Gated HRGC</td>
<td>1.013</td>
<td>0.191</td>
<td>1.204</td>
<td>4</td>
</tr>
<tr>
<td>Std Dev</td>
<td>0.764</td>
<td>0.329</td>
<td>0.895</td>
<td></td>
</tr>
<tr>
<td>Crossbuck Only</td>
<td>1.133</td>
<td>0.288</td>
<td>1.492</td>
<td>4</td>
</tr>
<tr>
<td>Std Dev</td>
<td>0.61</td>
<td>0.15</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Crossbuck with Yield</td>
<td>1.485</td>
<td>0.697</td>
<td>2.182</td>
<td>1</td>
</tr>
<tr>
<td>Crossbuck with Stop</td>
<td>1.414</td>
<td>0.648</td>
<td>2.062</td>
<td>2</td>
</tr>
<tr>
<td>Std Dev</td>
<td>0.093</td>
<td>0.101</td>
<td>0.008</td>
<td></td>
</tr>
</tbody>
</table>

Summary of Statistical Test Results for TCDs and at Accident Locations

As described in the methodology, we used a set of paired Welch’s t-tests to verify the statistical significance of the behavior score comparisons (Table 9).

Most t-tests failed to reject the null hypothesis, meaning that the average values shown were close enough for the means to be equal. The only pairs where the null hypothesis was rejected (results were statistically different) were comparisons between other TCDs and passive HRGCs with Stop signs (bold rows in Table 9).
Table 9 - Welch's T-test for comparison of TCD conditions

<table>
<thead>
<tr>
<th>Comparison Pair*</th>
<th>Scan Score</th>
<th>Speed Score</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>ES</td>
<td>p</td>
</tr>
<tr>
<td>Gates-Lights</td>
<td>0.7839</td>
<td>0.04</td>
<td>0.8003</td>
</tr>
<tr>
<td>Gates - Passive</td>
<td>0.5942</td>
<td>0.11</td>
<td>0.1694</td>
</tr>
<tr>
<td>Gates - Yield</td>
<td>0.9146</td>
<td>0.04</td>
<td>0.1421</td>
</tr>
<tr>
<td>Gates--Cross</td>
<td>0.9126</td>
<td>0.03</td>
<td>0.5411</td>
</tr>
<tr>
<td>Gates-Overall</td>
<td>0.8122</td>
<td>0.02</td>
<td>0.6705</td>
</tr>
<tr>
<td><strong>Gates-Stop</strong></td>
<td>0.0000</td>
<td><strong>1.10</strong></td>
<td>0.0000</td>
</tr>
<tr>
<td>Lights-Passive</td>
<td>0.7766</td>
<td>0.07</td>
<td>0.3065</td>
</tr>
<tr>
<td>Lights - Yield</td>
<td>0.4633</td>
<td>0.35</td>
<td>0.1658</td>
</tr>
<tr>
<td>Lights-Cross</td>
<td>0.9814</td>
<td>0.02</td>
<td>0.7056</td>
</tr>
<tr>
<td>Lights-Overall</td>
<td>0.8880</td>
<td>0.02</td>
<td>0.9853</td>
</tr>
<tr>
<td><strong>Lights-Stop</strong></td>
<td>0.0000</td>
<td><strong>1.10</strong></td>
<td>0.0000</td>
</tr>
<tr>
<td>Passive - Yield</td>
<td>0.5647</td>
<td>0.26</td>
<td>0.3515</td>
</tr>
<tr>
<td>Passive - Cross</td>
<td>0.7731</td>
<td>0.08</td>
<td>0.5961</td>
</tr>
<tr>
<td>Passive-Overall</td>
<td>0.6622</td>
<td>0.09</td>
<td>0.2241</td>
</tr>
<tr>
<td><strong>Passive - Stop</strong></td>
<td>0.0082</td>
<td>0.69</td>
<td>0.0000</td>
</tr>
<tr>
<td>Yield- Cross</td>
<td>0.4663</td>
<td>0.34</td>
<td>0.2320</td>
</tr>
<tr>
<td>Yield - Overall</td>
<td>0.4242</td>
<td>0.38</td>
<td>0.1578</td>
</tr>
<tr>
<td><strong>Yield - Stop</strong></td>
<td>0.9099</td>
<td>0.09</td>
<td>0.0060</td>
</tr>
<tr>
<td>Cross - Stop</td>
<td>0.0004</td>
<td>0.92</td>
<td>0.0000</td>
</tr>
<tr>
<td>Cross-Overall</td>
<td>0.9853</td>
<td>0.01</td>
<td>0.7054</td>
</tr>
<tr>
<td><strong>Overall – Stop</strong></td>
<td>0.0000</td>
<td>1.02</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

*Definitions: Gates – active HRGCs with gates; Lights – active HRGCs with lights, but no gates; Passive – Passive HRGCs, excluding those with stop signs; Yield – passive HRGCs with Yield signs; Cross – passive HRGCs with a crossbuck only; Stop – passive HRGCs with a stop sign; Overall – All HRGC locations together; p – probability that both means in the pair are the same; ES – Effect Size

Similar tests with the accident locations also failed to reject the null hypothesis between TCDs and between accident locations and the general population of HRGCs (Figure 10). There were no accident locations in the NDS data set for HRGCs with active lights but no gates and there was only one accident location with a yield sign, so the t-test could not be calculated.
Table 10 - Welch's t-test results for comparing accident to non-accident locations

<table>
<thead>
<tr>
<th>Comparison Pair*</th>
<th>Scan p</th>
<th>Score ES</th>
<th>Speed p</th>
<th>Score ES</th>
<th>Total p</th>
<th>Score ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall - All Locations</td>
<td>0.8843</td>
<td>0.11</td>
<td>0.6268</td>
<td>0.19</td>
<td>0.9719</td>
<td>0.02</td>
</tr>
<tr>
<td>Gates &amp; Lights</td>
<td>0.7473</td>
<td>0.64</td>
<td>0.8158</td>
<td>0.18</td>
<td>0.7192</td>
<td>0.51</td>
</tr>
<tr>
<td>Crossbuck only</td>
<td>0.8623</td>
<td>0.22</td>
<td>0.6100</td>
<td>0.21</td>
<td>0.8622</td>
<td>0.16</td>
</tr>
<tr>
<td>Crossbuck &amp; Stop Sign</td>
<td>0.9300</td>
<td>0.29</td>
<td>0.3849</td>
<td>0.72</td>
<td>0.3694</td>
<td>0.19</td>
</tr>
</tbody>
</table>

**Trending Parameter Analysis**

The main objective of this research program was comparison of accident/non-accident locations and the analysis of behavioral differences at different TCDs. However, we also performed preliminary trending analysis investigating the impacts of environmental factors and parameters on driver behavior scores, concentrating on factors that previous research has determined to impact the accident risk at HRGCs. To date, these analyses have been exploratory in nature and as such are not considered statistically valid.

We divided the analysis to two categories; 1) factors that are “crossing-specific” (Category 1), such as AADT, train volumes and train and highway speeds, and 2) factors that are ”traversal specific” (Category 2), such as weather, time of day, and driver demographics (age and gender).

**Category 1 – Crossing Specific Parameters**

Category 1 analysis concentrated on investigating whether differences in specific parameters that vary between HRGCs, but remain fixed within a single HRGC, have an effect on the mean driver behavior scores. Analysis in Category 1 were based on single factor analysis only, so the exclusion of other factors limits the robustness of the analysis. This could be improved through multivariate or clustering analysis in the future.
1.1.1.1 *Highway traffic levels (AADT)*

Figure 9 shows the average visual scan, speed, and total driver behavior scores versus Average Annual Daily Traffic (AADT) data, obtained from the FRA inventory database. Trendlines for both speed and visual scanning scores show slight decline, as traffic levels increase.

![Graph showing behavior scores versus AADT](image)

**Figure 10 – Visual Scan and Speed Scores by Annual Average Daily Traffic (Sample Size Per Category on the Top)**
1.1.1.2 Train traffic volumes

The Trains per Day (TPD) value is the sum of the day and night thru trains from the FRA inventory data base. Figure 10 presents the mean driver behavior scores. Note that there is a gap in the data between 25 and 50 trains per day, as no HRGCs had TPD values in that range. There were nearly 4,200 traversals at HRGCs with no thru trains reported in the FRA Inventory, reducing the sample size for the analysis. The analysis reveal increase in both speed and visual scan score as TPD value increases, although the latter takes place in a slower rate.

![Graph showing behavior score vs number of trains per day]

Figure 11 – Scan, Speed and Total Scores by Trains Per Day (Sample Size Per Category on the Top).

1.1.1.3 Highway Speed Limit

Figure 11 presents the mean behavior scores based on highway speed limit. There is a decrease in the scanning and speed scores as the highway speed increases. However, the data on the low end is limited, as the 10 mph and 15 mph analysis are based on less than 50 traversals each from a single HRGC. More than 1,000 data records were excluded, as they did not have a highway speed value posted in the FRA data base.
Figure 12 - Behavior Scores by Posted Highway Speed (Sample Size Per Category on the Top)
1.1.1.4 Train speed

Figure 12 compares the driver behavior scores with the train speed given in the FRA inventory. The scanning score appears to be consistent across the range of train speeds, but the speed score increases with higher train speed. A total of 107 traversals were omitted at these HRGCs. Note the jump in speed from 60 mph to 79 mph, as our data set had no records with 65 or 70 mph posted speeds.

Figure 13 - Behavior Scores by Train Speed (Sample Size Per Category on the Top)

Category 2 – Traversal Specific Parameters

Analysis on Category 2 parameters were conducted under a parallel project funded by the National University Rail Center (NURail) and concentrated on environmental factors (weather and time of day) and driver demographics (gender and age). Unlike in Category 1, these parameters do not remain fixed within a specific HRGC, but may vary between traversals based on prevailing conditions. The following is a summary of the findings from the analysis, including the results of the statistical tests. A more detailed description of the analysis is available in a report by Salim [18].
1.1.1.5 Weather conditions

Figure 13 shows the sample size and drivers’ mean behavior scores based on weather conditions. The data indicates that drivers received the lowest behavior scores in fog followed by rain, cloudy, clear and snow conditions. However, no conclusions could be made about drivers’ behavior in fog due to small sample size.
1.1.1.6 Time of Day

Based on previous studies, nighttime driving has been associated with more HRGCs accidents [1]. Figure 14 indicates sample size and average driver behaviors scores based on time of day. The figure reveals that drivers consistently received lower behavior scores during the night time traversals.

![Figure 15 - Drivers’ Mean Behavior Score Based on Time of Day](image)
1.1.1.7 Gender and Age

While driver demographics (gender and age) do not change between traversals, previous studies reported that male and younger drivers were involved in more HRGCs accidents compared to female and middle-aged drivers [1]. The results of our study, however, did not show any significant difference between the average driver behavior scores for male and female drivers (Figure 15).

![Graph showing average behavior scores by gender](image)

**Figure 16 – Mean Driver Behavior Scores by Gender**

Driver demographic data were further grouped into three age categories to compare behavior scores between different age groups of male and female drivers; younger adults (16-34 years old), middle-aged adults (35-54 years old) and older adults (55 years old or higher). Figure 16 shows the sample size and drivers’ average behavior scores based on gender and age groups. The data indicates that the difference in average behavior scores of male and female drivers in younger adults (1.4 versus 1.38) and older adults (1.39 versus 1.38) categories are negligible. The only noticeable difference was among middle aged drivers, where female drivers received higher behavior scores compared to male drivers in the same age category (1.44 versus 1.34). Note that 139 records did not have a valid age record.
Summary of Statistical Test Results for Weather, Time of day, Gender and Age

Table 11 shows the results of the t-tests conducted to compare parameter pairs under Category 2. Overall, several pairs (in bold letters) were found to have statistically significant differences in driver behavior score. From weather conditions, driver behavior under snowy conditions showed significant difference to most other conditions (clear, cloud, and rain). Perhaps the most consistent difference was shown between day time and night time traversals, with day time traversals witnessing significantly higher mean scores.
Table 11 - Summary of Statistical Test Results Based on Weather and Time of Day

<table>
<thead>
<tr>
<th>Condition Pair</th>
<th>Deg. Freedom</th>
<th>t Stat</th>
<th>P(T&lt;=t) two-tail</th>
<th>t Critical two-tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEATHER</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clear vs. Cloudy</td>
<td>233</td>
<td>1.09</td>
<td>0.28</td>
<td>1.97</td>
</tr>
<tr>
<td>Clear vs. Rain</td>
<td>472</td>
<td>2.70</td>
<td>0.01</td>
<td>1.97</td>
</tr>
<tr>
<td>Clear vs. Snow</td>
<td>92</td>
<td>-2.29</td>
<td>0.02</td>
<td>1.99</td>
</tr>
<tr>
<td>Cloudy vs. Rain</td>
<td>432</td>
<td>0.66</td>
<td>0.51</td>
<td>1.97</td>
</tr>
<tr>
<td>Cloudy vs. Snow</td>
<td>166</td>
<td>-2.54</td>
<td>0.01</td>
<td>1.97</td>
</tr>
<tr>
<td>Rain vs. Snow</td>
<td>127</td>
<td>-3.22</td>
<td>0.002</td>
<td>1.98</td>
</tr>
<tr>
<td>TIME OF DAY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day vs. Night</td>
<td>2267</td>
<td>6.82</td>
<td>1.1E-11</td>
<td>1.65</td>
</tr>
</tbody>
</table>

We found no statistically significant difference between male and female drivers or between different age groups. However, when gender and age parameters were combined (Table 12), males in the 35-54 years old category had significantly lower average scores than females in same age group.

Table 12 - Summary of Statistical Test Results Based on Gender and Age

<table>
<thead>
<tr>
<th>Condition Pair</th>
<th>Age Group</th>
<th>Deg. Freedom</th>
<th>t Stat</th>
<th>P(T&lt;=t) two-tail</th>
<th>t Critical two-tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male vs. Female</td>
<td>16-34</td>
<td>4330</td>
<td>0.79</td>
<td>0.43</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>35-54</td>
<td>1673</td>
<td>-2.32</td>
<td>0.02</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>55+</td>
<td>2662</td>
<td>0.41</td>
<td>0.68</td>
<td>1.96</td>
</tr>
</tbody>
</table>
• **Discussion and Limitations of Phase 1**

We used a three-point driver behavior score methodology developed at Michigan Tech and over 9,000 records of NDS data from the SHRP2 program to quantitatively evaluate driver behavior at HRGCs. Both the use of NDS data for the analysis and scoring methodology were novel and based on drivers visual scanning and speed adjustment while approaching HRGCs. A higher behavior score was considered an indication of a more defensive driving behavior. We summarize our key findings from the Phase 1 in the following sections.

**TCD Analysis**

The results from the TCD analysis revealed no statistically significant difference in driver behavior at HRGCs with various types of TCDs, except for HRGCs with passive warnings that included stop signs. In general, the total driver behavior scores in HRGCs with passive warning devices aligned fairly closely with those equipped with active warning devices, even if the visual scanning and speed adjustment components of the score were investigated separately. The average scanning scores of 1.13 suggest that drivers are scanning in both directions during approximately 60% of the traversals, while the speed score of approximately 0.284 suggests that drivers are preparing to stop at 30% of the HRGCs. A closer look at the actual scores revealed that in approximately 34% of the traversals drivers look both ways, 48% look one way, and 18% fail to look at all. These values were consistent across all types of TCDs, except HRGCs with stop signs. Scores for HRGCs with active TCDs showed speed reduction approximately 24% of the time, while the corresponding percentage at passive HRGCs with crossbucks warning only was lower (20%). The results were slightly better for yield controlled HRGCs, with 32% of drivers preparing to stop, but they were based on only seven HRGCs. Overall, we believe that the equal or even lower scores of visual scanning/speed adjustment at passive HRGCs is concerning. While locations with active TCDs provide visual warning of approaching trains to drivers, their only defense at passive HRGCs is through individual recognition of potential danger. The majority of drivers that do not prepare for potential train arrival place themselves at risk during every traversal, providing one explanation for higher accident rates witnessed over time at passive HRGCs.

As mentioned earlier, the only statistically significant exception were higher mean scores at passive HRGSs equipped with stop signs. The overall difference was mainly caused by better performance in the speed reduction category, as the percentage of drivers preparing to stop increased to above 70% when a stop sign was present. This percentage might be even higher, as in some traversals a low entry speed, combined with failure to come to a complete stop, led to an insufficient speed reduction to score a point. We interpret this result as confirmation that drivers express a higher level of defensive driving at HRGCs with stop signs and as such, the requirements in the MUTCD to increase their usage [19] seem warranted. On the other hand, it remains to be seen how permanent the improvement in driver behavior is, once drivers start to encounter them more frequently.

**Accident Analysis**

Fifteen HRGCs in our NDS database had witnessed accidents between 2000 and 2010, reducing the overall sample size available for the analysis. We found that mean driver behavior scores at accident locations were similar to corresponding values in the full sample. The only notable difference was the higher mean score at the one (the only) passive HRGC with yield sign, but this
also highlights the shortcoming of the analysis, i.e. the small number of HRGCs in the accident dataset. Overall, the small sample size makes the results questionable and no conclusions could be made to differentiate driver behavior between HRGCs with and without accidents.

Trending Analysis

As noted before, the initial trending analyses were based on looking at parameters individually and their impact on the behavior score. The analyses were broken down to two categories: Category 1 that concentrated on “crossing specific” parameters, and Category 2 that concentrated on “traversal specific” parameters. For Category 1, the analyses were exploratory in nature and not tested for statistical significance.

**Category 1 – Crossing Specific Parameters**

The trending analysis on crossing specific parameters showed several linear trends across the various values. While the statistical significance was not tested, the results can be considered at least as valuable observations that include the following:

- The preliminary trend analysis (Figure 12) revealed a moderate decrease in both the scanning and speed scores as AADT increases. One explanation may relate to the fact that as AADT increases the use of active TCDs can also be expected to increase. The added reliance on the active TCDs is likely to reduce the mean driver behavior scores. In addition, with higher AADT drivers are more likely to follow the traffic flow when traversing HRGC.

- Scanning behavior appears to increase slightly as the number of trains per day (TPD) increases (Figure 13). Interestingly, the scan scores for HRGCs with no reported thru trains are very similar to the rest of the HRGCs in our database. Speed scores do appear to increase with train traffic, suggesting that drivers are more likely to reduce their speed with more frequent train traffic.

- There seems to be a clear trend for both lower scanning and speed behavior scores as highway speed limit increases (Figure 14). Just like with AADT, we believe this decline to be related to increased presence of active TCDs on roadways with higher speed limits. However, the results could also be interpreted to suggest that at higher speeds driver attention is more focused on the features within the roadway, leaving less attention for potential trains.

- The scanning behavior does not seem to be affected by the train speed at HRGCs (Figure 15). However, the speed score shows increase as train speed increases, suggesting that drivers are more aware of the potential danger from higher speed trains.

**Category 2 – Traversal Specific Parameters**

While the Category 2 analysis did not reveal any major trends between the scenarios, we found statistically significant differences between several condition pairs tested.

- We found that under different weather conditions, drivers received the highest behavior scores in snow and lowest in rainy conditions (fog was even lower, but the sample size was too small for statistical analysis). Since low driver behavior score was indicative of less defensive driving, the results support the findings from previous studies for increased risk
in rainy conditions. However, they do not support the previous finding for increased accident risks under snowy conditions.

- The results were more consistent on differences in driver behavior based on **time of day**. We found that all drivers from all gender/age groups received significantly lower behavior scores during the night compared with day. This outcome supports previous studies that have revealed poor visibility conditions to have negative impact on driver behavior, thus increasing accident risk at HRGCs [15]. Based on our study results, we suspect that poor visibility at night time lead to higher concentration on the road ahead and as such to lower driver behavior score.

- We did not find any significant difference in average behavior scores of **male and female drivers**. Based on driver demographic analysis, the only statistically significant difference was between **middle-aged** female drivers (35-54 years old) when compared to male drivers in the same age category. This differs from earlier studies that concentrated on accidents at HRGCs and found younger male drivers to be at higher risk.
2. Phase 2 – Data Sources and Methodology for Driver Simulator Study

We selected two HRGC locations (crossing ID’s: 521090P, 621549W) from the RID database for replication in the driver simulator. These crossings were selected based on the number of traversals available in the NDS data and the expected level of difficulty for modeling the specific HRGC environment in the simulator. The HRGCs were virtually modeled in collaboration with the National Advanced Driving Simulator (NADS) team at Iowa State University, using NADS MiniSimTM software (Appendix B).

The first HRGC (scenario A) featured active TCDs (gates and flashing lights) and the second HRGC (scenario C) featured cantilevered flashing lights only as the active TCD. The simulator studies were conducted in two batches, but the main analysis combined studies one and two into a single “SIM” dataset.

Table 13 presents a summary of parameters for both NDS and simulator analysis for ease of comparison. Driver behavior was quantified using the same driver behavior score as in the NDS data analysis. Since we had access to the un-blurred video of the participants face in the SIM dataset, visual scanning behavior was coded manually, as opposed to automatic classification based on estimated head tracking data.

<table>
<thead>
<tr>
<th>92.</th>
<th>93. NDS</th>
<th>94. SIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>95. Speed limit</td>
<td>96. 45 mph</td>
<td>97. Study 1 - no limit</td>
</tr>
<tr>
<td>98.</td>
<td>99.</td>
<td>100. Study 2 - 45 mph</td>
</tr>
<tr>
<td>101. # of Drivers</td>
<td>102. 24</td>
<td>103. 48 (in analysis)</td>
</tr>
<tr>
<td>104. # of Traversals</td>
<td>105. 284 (total)</td>
<td>108. 256 (total)</td>
</tr>
<tr>
<td></td>
<td>107. Scenario C (lights only) - 233</td>
<td>110. Scenario C (lights only) - 132</td>
</tr>
<tr>
<td>111. Advanced Warning Distance</td>
<td>112. Determined from videos and Google Earth</td>
<td>113. (same as NDS)</td>
</tr>
<tr>
<td>117. Source of speed analysis data</td>
<td>118. NDS vehicle data</td>
<td>119. SIM vehicle data</td>
</tr>
</tbody>
</table>

A total of 51 traversals were included in the NDS dataset for scenario A (gates/lights), and 233 traversals were included in the NDS dataset for scenario C (lights only). This dataset included 24
unique drivers ($M_{age} = 41.8$, $SD_{age} = 23.5$, 13 male, 11 female). For the simulator study, a total of 54 participants ($M_{age} = 20.12$, $SD_{age} = 1.71$; 40 male, 14 female) were recruited from Michigan Tech’s undergraduate psychology courses. The participant sample had an average of 4.83 years ($SD_{years} = 1.71$) of driving experience.

We collected data in two closely aligned studies. Study 1 included 20 participants, and Study 2 included the remaining 34. Participants were compensated with course credit in exchange for their time. They were not informed about the specific goals of the study beforehand, as the recruitment advertisements only mentioned that the experiment was about “driver behavior in a medium fidelity simulator”.

The final dataset for Study 1 analysis included 18 participants, and for Study 2 included 30. The remaining six participants were dropped from later analysis due to missing data caused by experimenter error or technological difficulties. The final SIM dataset for the analysis included 124 simulated traversals at the crossing with lights/gates (scenario A), and 132 traversals at the crossing protected by lights only (scenario C).

All experimental stimuli and protocol were identical in both simulator studies. The only difference were the instructions for participants. In Study 1, participants were instructed to ignore the simulated speedometer and only use speed perception cues from the simulated environment when determining the appropriate speed to drive. This was due to the pilot participant feedback that suggested a mismatch between the visual cues of motion and the speedometer in the simulator. In Study 2, participants were instructed to obey a 45-mph speed limit throughout the entire experimental session.

### 2.1.1 Experimental Design

We used two existing HRGCs in rural setting (included in our NDS dataset) to develop corresponding simulated scenarios (Figure 17). Each scenario was included in a loop that participants drove around three times in succession. Each lap included two HRGCs, resulting in six traversals per scenario for each participant. Study 1 included a train present event on the final ($6^{th}$) traversal for both scenarios, but otherwise both studies had identical setting (excluding the speedometer instructions). As described in previous section, the crossings included HRGC with lights and gates (Figure 18), and HRGC with flashing lights in cantilevered support (Figure 19). The order of scenario exposure was counterbalanced across participants, meaning half the participants experienced scenario A first, and the other half experienced scenario C first. Each scenario was around 20 minutes in length, depending on the speed of the participant.
Figure 18 – Track Designs for Scenario A (left) and C (right). Red Lines Indicate Railroads That Intersect with Highways (Grey).

Figure 19 – Simulated advance warning (left) and gated crossing (right) in scenario A

Figure 20 – Simulated Advance Warning in Scenario C (Left) Compared to Reference NDS Video (Right)

After completing both scenarios, participants filled out the following post-experiment survey:

1. How many years have you been driving for?
2. How realistic were the scenarios you drove in?
3. How different is your real-world driving behaviors from your simulated driving behavior?
4. How many times do you encounter a train crossing per month (on average)?
5. How many times have you had to come to a stop for a train at a road crossing in your lifetime?
6. What would/should you do when encountering this sign? (the question would present a picture of each of the different HRGC TCDs featured in the experiment).
7. How did your behavior at railroad crossings change over the course of this experiment?
8. How noticeable was the head-tracking device and cap? Do you think it had any impact on your behavior?

### 2.2 Results and Data Analysis

First, we conducted a 2x3 analysis of variance (ANOVA, Table 14) to detect significant differences between simulated and naturalistic data sources (SIM vs. NDS) and TCD types (Gate/Lights vs. Lights only) on driver behavior scores. Groups were organized by data source (simulated vs. naturalistic data sets) and TCD types (Gate/Lights (A) vs. Lights only (C) scenario). Table 14 reveals that the P value for each is below 0.05, suggesting that significant main effects on driver behavior scores exist for both TCD types and data source, as well as for their interaction.

**Table 14 – 2x3 Analysis of Variance**

<table>
<thead>
<tr>
<th>Category</th>
<th>DF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>P&lt;0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Source</td>
<td>2</td>
<td>5.04</td>
<td>2.519</td>
<td>4.78</td>
<td>0.008*</td>
</tr>
<tr>
<td>Scenario (TCD Type)</td>
<td>1</td>
<td>2.85</td>
<td>2.848</td>
<td>5.41</td>
<td>0.02*</td>
</tr>
<tr>
<td>Source: Scenario</td>
<td>2</td>
<td>8.47</td>
<td>4.237</td>
<td>8.05</td>
<td>0.0003</td>
</tr>
<tr>
<td>Residuals</td>
<td>53</td>
<td>280.9</td>
<td>0.526</td>
<td>148.</td>
<td>149.</td>
</tr>
</tbody>
</table>

We conducted an independent sample T-test to compare total driver behavior scores between the SIM and NDS data sets. Results suggested driver behavior scores were significantly higher in the NDS dataset than the SIM dataset ($t(533) = 2.99, p = .003$). Figure 20 depicts mean behavior score across data sets. The error bars show standard error of the mean. The left graph’s y axis is scaled to show the full 3-point range of the behavior score. The right graph depicts the same data, only zoomed in for easier visual comparison. A similar method is used for the following graphs as well.
We conducted a second independent sample T-test to compare the behavior scores between the two types of TCDs (scenario A vs. C). Results suggested driver behavior scores were not significantly different between the scenarios across both NDS and SIM data sets ($t(310) = 1.17$, $p = .24$). Figure 21 depicts mean behavior scores for both scenarios (types of TCD).

A significant interaction effect suggests the effect of TCD type is larger in the NDS dataset than for the SIM dataset (Figure 22). When only considering SIM data, there is no statistical difference between scenario A and C. However, when only considering NDS data, driver behavior scores were higher in response to scenario A (gates/lights) than for scenario C (lights only). For the
interaction plot in Figure 23, the small dots represent individual data points with artificial jitter (right). A significant interaction difference is indicated by the slopes between the two data sources when comparing responses to different TCD types.
To further explore the impacts of the different instructions between Study 1 and Study 2, the SIM dataset was split between the respective studies (Figure 24). A post-hoc Tukey Honest Significant Difference (HSD) test for multiple comparisons was run to investigate the differences between each subgroup (Table 15). The post-hoc Tukey HSD runs multiple unpaired t-tests simultaneously while controlling for the additional family-wise type 1 error rate (false positives due to multiple comparisons). The results suggest that NDS behavior scores were significantly higher than SIM scores in Study 1 while the difference between NDS and Study 2 scores fell just short of the significance criteria ($p = .056$). Figure 25 depicts the means and standard errors for each subgroup to visualize the interaction between data source and TCD type.
Table 15 - Post-hoc Tukey HSD Analysis

<table>
<thead>
<tr>
<th>Subgroups</th>
<th>Diff</th>
<th>lwr</th>
<th>upr</th>
<th>p adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIM1-NDS</td>
<td>-0.223</td>
<td>-0.425</td>
<td>-0.021</td>
<td>0.026*</td>
</tr>
<tr>
<td>SIM2-NDS</td>
<td>-0.168</td>
<td>-0.341</td>
<td>0.003</td>
<td>0.056</td>
</tr>
<tr>
<td>SIM2-SIM1</td>
<td>0.054</td>
<td>-0.167</td>
<td>0.277</td>
<td>0.831</td>
</tr>
</tbody>
</table>

Figure 25 – Mean Behavior Score Grouped by NDS and SIM (Study 1 and Study 2) Data Sets

Figure 26 – Mean Behavior Score Grouped by NDS, Study 1, and Study 2 Data Sets.
2.2.1 Summary of Behavior Scores

Figure 26 is a stacked bar chart depicting the contribution of each behavior score component (visual scanning vs. speed reduction) to the total driver behavior score. The yellow portion represents the mean speed reduction score (total 1 point possible), and the grey portion represents the mean visual glance score (total 2 points possible). In both data sets (NDS and SIM), most of the scores received by the participants are from visual glances, and very few are from speed reduction. Only 6.3% (18/284) of traversals received a speed reduction point in the NDS dataset, and only 3.1% (8/256) of traversals received a speed reduction point in the SIM dataset (Table 13).

![Stacked Bar Chart Depicting the Contribution From Behavior Score Sub-Behaviors](image)

Table 16 provides the percentage distribution of the four possible driver behavior scores (0, 1, 2, 3). The score patterns are different for NDS and SIM datasets for Scenario A (gates/lights). In the NDS dataset, a great majority of drivers scored 1 or 2 points while less than half did in the SIM dataset.

We also observed a low number of drivers scoring a point from speed reduction in both data sets, 6.3% and 3.1%, respectively (Table 17). Hence, majority of points came from visual scanning behavior.
Table 16 - Percent of Total Behavior Score

<table>
<thead>
<tr>
<th>TCD type</th>
<th>Score bin</th>
<th>NDS count</th>
<th>NDS percentage</th>
<th>SIM Count</th>
<th>SIM percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate/Lights</td>
<td>0</td>
<td>9</td>
<td>17.65%</td>
<td>71</td>
<td>57.26%</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>22</td>
<td>43.14%</td>
<td>98</td>
<td>28.23%</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>20</td>
<td>39.22%</td>
<td>32</td>
<td>13.71%</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0</td>
<td>0%</td>
<td>1</td>
<td>0.81%</td>
</tr>
<tr>
<td>Lights Only</td>
<td>0</td>
<td>102</td>
<td>43.78%</td>
<td>64</td>
<td>48.48%</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>98</td>
<td>42.06%</td>
<td>54</td>
<td>40.91%</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>32</td>
<td>13.73%</td>
<td>12</td>
<td>9.09%</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>1</td>
<td>0.43%</td>
<td>2</td>
<td>1.52%</td>
</tr>
</tbody>
</table>

Table 17 - Count and Percent of Traversals With Speed Reduction During the Taversal

<table>
<thead>
<tr>
<th>Speed reduction point awarded</th>
<th>NDS</th>
<th>SIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>fail</td>
<td>266</td>
<td>248</td>
</tr>
<tr>
<td>pass</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Pass Percent</td>
<td>6.30%</td>
<td>3.10%</td>
</tr>
</tbody>
</table>

2.2.2 Summary of Post-Experiment Survey

The following highlights some of the key findings from the post-experiment survey.

- All but one participant reported that the driving simulator scenarios were realistic. A few participants suggested the low visual quality and the unnatural response of the vehicle as two potential features that could detract from the realism of the simulation.

- Roughly half of the participants reported that there was little to no difference between their driving style in the simulator and their driving style in the real world. Fourteen participants admitted that they drove less cautiously in the simulation due to the lack of consequences. Alternatively, seven participants admitted that they drove more cautiously in the simulation because it was an experiment where their behavior was being closely monitored.

- Participants were shown pictures of each of the TCDs and were asked to describe the appropriate driver response to each. All participants understood that it was their responsibility to observe and yield to oncoming trains at HRGCs with passive TCDs. Approximately half of the participants reported that they would approach active TCDs in the off position with caution by slowing down and scanning for trains. The other half suggested that active TCDs in the off position indicate that drivers do not need to slow down or look for the trains.
Almost half of the participants experience one to three HRGCs on a typical month of driving (Figure 27 – top) and almost half of the participants have come to a complete stop because of a train less than five times in their lifetime (Figure 27 – bottom).

![Pie Graphs of Exposure to Crossings and Trains in Real World Driving](image)

**Figure 28 – Pie Graphs of Exposure to Crossings and Trains in Real World Driving**

### 2.3 Discussion of Phase 2

We used two existing HRGCs included in the NDS dataset to develop simulated scenarios in the National Advanced Driving Simulator (NADS). We then recruited student drivers to investigate driver behavior at HRGCs in the simulated settings and to compare and contrast driver behavior between the NDS and simulated data sets. The evaluations used the same 3-point quantitative method as the NDS analysis. The following section briefly discusses the key findings from Phase 2 of the study.

- Post-experiment survey revealed that the simulator was considered fairly realistic and participants maintained their driving style. However, the higher operating speeds by Study
1 participants (and the post-experiment survey results) suggest that the speed perception in the simulator differs from that in the natural environment.

- Comparisons between SIM data from Study 1 and Study 2 highlight the effects of train presence and verbal speed instructions, as those were the only differences between the two studies. Participants drove around 80 mph in the simulator without the explicit verbal instructions to follow the 45mph speed limit. However, we found no statistical difference for mean behaviors scores between stimulation studies 1 and 2, which suggests participants approached HRGC with the same amount of caution regardless of average vehicle speed and the inclusion/exclusion of train present events.

Our main findings from comparing and contrasting the results of NDS and SIM data sets are as follows:

- Both data sets suggest that drivers’ behavior at the crossings have room for improvement from visual scanning and speed adjustment perspective. Mean driver behavior scores are low for both NDS and SIM datasets, 0.8 and 0.6 out of possible 3 total points, respectively. Especially, speed reduction/adjustment behavior is poor in both NDS and SIM datasets. Very few crossing events (26/540, or 4.8% of total) in either data set (NDS and SIM) were awarded a point for speed reduction using current criteria and very few drivers (4/540, or 0.7% of total) received the full 3 points for the driver behavior score.

- A closer look reveals the difference between the two data sets. Statistical tests (ANOVA and t-test) suggested statistically significant difference in the driver behavior scores between NDS and SIM data sets, but not between the different TCDs (Scenario A vs. C).

- The significant main effect for data source suggests participants in the simulation approached HRGCs less cautiously than participants in the NDS dataset. The significant main effect for TCD type suggests drivers approached HRGC with lights/gates more cautiously than HRGC with flashing lights only, but we only observed this trend in the NDS dataset in post-hoc subgroup analyses. Similarly, behavior scores were higher for scenario A (gates/lights) in the NDS data set, but the same trend was not observed in the SIM dataset.

- The post-experiment survey revealed participants’ mixed knowledge and the gap between their knowledge and behavior. Some participants drove in the simulator more cautiously than in their actual driving, while other participants drove less cautiously. Also, half of them reported they would approach active warnings in the off position with caution, but the other half reported they would not slow down or scan for trains because warning is off. All participants understood that it was their responsibility to search for and yield to oncoming trains for passive RR warnings, but (consistent with other driving behavior research) their subjective responses do not correspond with their actual behavior patterns.
• **Study Limitations**

Based on our knowledge, this study was the first attempt to use the SHRP2 NDS data for the driver behavior analysis at the HRGCs. We developed a novel “driver behavior score” methodology for the analysis and for the comparisons with simulator experiments conducted in Phase 2. We want to highlight the following limitations and challenges identified in our data and approach during the study.

- **Completeness and accuracy of NDS data**: Due to the difficulties with keeping the sensor arrays in 3,500 vehicles running all the time over an extended period of time, some of our NDS data records had missing or incomplete data. For example, some data records missed GPS data and numerous records didn’t have data for gas or brake pedal depressions. We have investigated methods for circumventing the missing data, and have developed some techniques to bring some of the excluded records back into our analysis.

- **Use of a 3-point score as the sole qualifier for the analysis**: For simplicity, our methodology relied on developing a single driver behavior score, based on two activities (head rotation and speed reduction) as a quantitative indicator for all behavior during the HRGC traversal. However, condensing a whole chain of events into a single score limits the possibility to investigate the impact of specific factors on driver behavior. It also combines large number of HRGCs with varying characteristic and excludes certain types of HRGCs from the analysis (such as HRGCs near highway intersections). An alternative way to use NDS data in HRGC safety analysis is concentrating on a single behavior, such as the location of speed reduction. Such targeted analysis might provide more granular data for parametric safety analysis.

- **Limited sample size**: While over 9,000 samples is significant amount of data for the analysis, its division into numerous subcategories reduces the sample size per category. Especially the low number of HRGCs with passive TCDs limits the comparative analysis. This is also true on the simulator side. Two HRGCs and 40+ candidates with limited background diversity are at a best “a good start” for statistical analysis.

- **Simulator perception**: Research has shown that the level of “realism” or the fidelity of the driving simulator scenario do not necessarily influence the study outcomes. However, we acknowledge that the perception about the actual risk may be different between the two and can affect the study outcomes.
The Phase 1 of this study we used over 9,000 individual traversals obtained from the SHRP 2 naturalistic driving study (NDS) data to evaluate driver behavior at HRGCs. We developed a three-point behavior score based on visual scanning for trains and vehicle speed reduction for the quantitative evaluation. Mean scores were used to generate comparisons and trends, based on selected parameters and conditions. The main comparisons included HRGCs divided by different TCDs and with or without an accident history. In Phase 2, we implemented selected HRGCs from NDS data set in a driver simulator. We used the same evaluation methodology to compare and contrast the simulator results with NDS ones.

While the NDS data analysis resulted in numerous interesting observations, they showed little statistical difference in driving behavior between any of the TCDs analyzed. The only exception was the significantly higher mean scores at passive HRGCs equipped with stop signs. The other consistent finding was the higher mean scores for traversals that took place during the day versus night time. We found no significant evidence of systematic driver behavior differences in most other categories tested, such as behavior between genders or between age groups.

It was evident from the NDS data (and driver behavior score) that most drivers were not scanning for trains, nor were they preparing to stop, even at crossings with passive warning devices where drivers must rely on their own observations for a safe passage. For example, in the driver simulator scenarios, less than 10 percent of all participants (NDS and simulator data combined) received a point for a proper speed adjustment. With this result, we plan to conduct further research on how to influence drivers to behave appropriately using additional effective but cost-efficient methods (e.g., in-vehicle alerts).

In Phase 2, we used driver simulator experiment and related comparisons to investigate similarities and differences in the driver behavior scores between simulated and natural settings. In general, similar trends could be observed in each data set, but there were more significant differences in scores between scenarios in NDS than in simulated setting, a finding that may be attributed to the relative perception of “safeness” of the simulated environment.

We remain convinced that better understanding driver behavior at HRGCs would be of value when predicting situations when drivers are less cautious and could be at risk of accidents. Despite the limitations and shortcomings of the current effort, we feel confident that NDS data can be beneficial in improving that understanding. We also found no reason to invalidate the use of driver simulators in the analysis. Instead, we believe that they allow researchers to quickly and efficiently analyze new methods to address driver behavior and potential new safety improvements at HRGCs.

The methodologies and data processes developed in this study to evaluate parameters, such as weather, demographics, time of day, etc. are by no means perfect, but they do offer an opportunity to try to quantify the performance of everyday drivers on a large scale. The limitations identified in the report should be addressed and our methodologies modified to improve the accuracy and credibility of the analysis. Instead of relying on a single three-point scale to evaluate the complete behavior, the next steps in the research may concentrate on improving the understanding of individual parameters, such as the exact locations where drivers adjust their speed when approaching HRGC. Research may also consider the use of multivariate analysis techniques to investigate which environmental variables, or groups of variables, have the most significant effects.
on HRGC behavior. We also believe that value can be found in harnessing machine learning and artificial intelligence applications for predicting driver behavior and their effects on perceived risks.
• References


[18] A. Salim, “Evaluation of Driver Behavior at Highway-Railroad Grade Crossings Based On Environmental Conditions And Driver Demographics”, MS Report, Michigan Technological University, 2018

Appendix A. - Percent Behavior Scores

Table 18 presents an analysis showing the percentage of drivers who received each level of behavior score. The table shows total score as well as scan and speed scores separately. Figures 28-30 show the same comparisons in a graphic format.

Table 18 - Behavior Score by Percent

<table>
<thead>
<tr>
<th>150. Scan Score</th>
<th>151. # Traversals</th>
<th>152. Percent</th>
<th>153. Speed Score</th>
<th>154. # Scores</th>
<th>155. Percent</th>
<th>156. Total Score</th>
<th>157. # Scores</th>
<th>158. Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>160.</td>
<td>161.</td>
<td>162.</td>
<td>163.</td>
<td>164.</td>
<td>165.</td>
<td>166.</td>
<td>167.</td>
<td></td>
</tr>
<tr>
<td>195.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>196.9128</td>
<td>197.</td>
<td>198.</td>
<td>199.</td>
<td>200.</td>
<td>201.0</td>
<td>202.134</td>
<td>203.148</td>
<td></td>
</tr>
<tr>
<td>205.</td>
<td>206.</td>
<td>207.</td>
<td>208.</td>
<td>209.</td>
<td>210.0</td>
<td>211.</td>
<td>212.0</td>
<td></td>
</tr>
<tr>
<td>241.219</td>
<td>242.</td>
<td>243.</td>
<td>244.</td>
<td>245.</td>
<td>246.0</td>
<td>247.31</td>
<td>248.142</td>
<td></td>
</tr>
<tr>
<td>250.</td>
<td>251.</td>
<td>252.</td>
<td>253.</td>
<td>254.</td>
<td>255.</td>
<td>256.</td>
<td>257.</td>
<td></td>
</tr>
<tr>
<td>285. Total</td>
<td>286.535</td>
<td>287.</td>
<td>288.</td>
<td>289.</td>
<td>290.</td>
<td>291.0</td>
<td>292.80</td>
<td>293.150</td>
</tr>
<tr>
<td>295.</td>
<td>296.</td>
<td>297.</td>
<td>298.</td>
<td>299.</td>
<td>300.</td>
<td>301.</td>
<td>302.</td>
<td></td>
</tr>
<tr>
<td>330. Total</td>
<td>331.1387</td>
<td>332.</td>
<td>333.</td>
<td>334.</td>
<td>335.</td>
<td>336.0</td>
<td>337.202</td>
<td>338.146</td>
</tr>
<tr>
<td>348.2</td>
<td>349.2353</td>
<td>350.33.</td>
<td>351.1</td>
<td>352.163</td>
<td>353.23.3</td>
<td>354.3</td>
<td>355.576</td>
<td>356.8.2</td>
</tr>
<tr>
<td>Score</td>
<td>All Traversals</td>
<td>Crossbuck with Yield</td>
<td>Crossbuck Only</td>
<td>Lights, No Gates</td>
<td>Gates</td>
<td>Crossbuck with Stop</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>----------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-------</td>
<td>---------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>357.1</td>
<td>358.3344</td>
<td>359.47.9</td>
<td>360.0</td>
<td>361.535.7</td>
<td>362.76.7</td>
<td>363.2</td>
<td>364.257.7</td>
<td>365.36.9</td>
</tr>
<tr>
<td>366.0</td>
<td>367.1290</td>
<td>368.18.5</td>
<td>369.0</td>
<td>370.0</td>
<td>371.0</td>
<td>372.1</td>
<td>373.279.8</td>
<td>374.40.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>376.6987</strong></td>
<td><strong>377.0</strong></td>
<td><strong>378.0</strong></td>
<td><strong>379.0</strong></td>
<td><strong>380.0</strong></td>
<td><strong>381.0</strong></td>
<td><strong>382.103.6</strong></td>
<td><strong>383.14.8</strong></td>
</tr>
<tr>
<td><strong>Crossbuck with Stop</strong> &amp; <strong>Total</strong></td>
<td>385.0</td>
<td>386.0</td>
<td>387.0</td>
<td>388.0</td>
<td>389.0</td>
<td>390.0</td>
<td>391.0</td>
<td>392.0</td>
</tr>
<tr>
<td>393.2</td>
<td>394.73</td>
<td>395.44.8</td>
<td>396.1</td>
<td>397.125</td>
<td>398.72.3</td>
<td>399.3</td>
<td>400.52</td>
<td>401.31.9</td>
</tr>
<tr>
<td>402.1</td>
<td>403.75</td>
<td>404.46.0</td>
<td>405.0</td>
<td>406.48</td>
<td>407.27.7</td>
<td>408.2</td>
<td>409.73</td>
<td>410.44.8</td>
</tr>
<tr>
<td>411.0</td>
<td>412.15</td>
<td>413.9.2</td>
<td>414.0</td>
<td>415.0</td>
<td>416.0</td>
<td>417.1</td>
<td>418.37</td>
<td>419.22.7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>421.163</strong></td>
<td>422.0</td>
<td>423.0</td>
<td>424.0</td>
<td>425.0</td>
<td>426.0</td>
<td>427.1</td>
<td>428.0.6</td>
</tr>
</tbody>
</table>

**Figure 29 - Percent of Traversals with the Given Total Score, by TCD**
Figure 30 - Percentage of Traversals with the Given Scan Score by TCD

Figure 31 - Percentage of Traversals with the Given Speed Score by TCD
Appendix B. Driving Simulator Description

The driving simulator runs the National Advanced Driving Simulator (NADS) MiniSim version 2.2 software. The hardware (Figure 31) includes a single computer, running Microsoft Windows 7 Pro on an Intel Core i7 processor, 3.07 GHz and 12 GB of RAM, and relays sound through a 2.1 audio system. Three Panasonic TH42PH2014 42" plasma displays with a 1280x800 resolution each allow for a 130 degree field of view in front of the seated participant. The center monitor is 28 inches from the center of the steering wheel and the left and right monitors are 37 inches from the center of the steering wheel. The MiniSim also includes a real steering wheel, adjustable car seat, gear-shift, and gas and brake pedals, as well as a Toshiba Ltd. WXGA TFT LCD monitor with a 1280x800 resolution to display the speedometer, etc. Environmental sound effects are also played through two embedded speakers. These sounds included engine noise, brake screech, turn indicators, collisions, auditory alerts, etc. In the present experiment, all participants experienced the same pre-defined route and properties for the driving task.

Figure 32 – NADS MiniSim Driving Simulator