

Guidelines for safer pedestrian crossings: Understanding the factors that positively influence vehicle yielding to pedestrians at unsignalized intersections

Raphael Stern, Principal Investigator
Civil, Environmental, and Geo- Engineering
University of Minnesota

JUNE 2023

Research Project
Final Report 2023-24



To request this document in an alternative format, such as braille or large print, call [651-366-4718](tel:651-366-4718) or [1-800-657-3774](tel:1-800-657-3774) (Greater Minnesota) or email your request to ADArequest.dot@state.mn.us. Please request at least one week in advance.

Technical Report Documentation Page

1. Report No. MN 2023-24		2.		3. Recipients Accession No.	
4. Title and Subtitle Guidelines for safer pedestrian crossings: Understanding the factors that positively influence vehicle yielding to pedestrians at unsignalized intersections				5. Report Date June 2023	
				6.	
7. Author(s) Raphael Stern, Tianyi Li				8. Performing Organization Report No.	
9. Performing Organization Name and Address Department of Civil, Environmental, and Geo- Engineering University of Minnesota 500 Pillsbury Dr SE Minneapolis, MN 55455				10. Project/Task/Work Unit No. CTS #2021008	
				11. Contract (C) or Grant (G) No. (c) 1036210	
12. Sponsoring Organization Name and Address Minnesota Department of Transportation Office of Research & Innovation 395 John Ireland Boulevard, MS 330 St. Paul, Minnesota 55155-1899				13. Type of Report and Period Covered Final Report	
				14. Sponsoring Agency Code	
15. Supplementary Notes http://mdl.mndot.gov/					
16. Abstract (Limit: 250 words) Many factors influence an individual driver's decision to yield or not yield to individual pedestrians attempting to cross the road at an unsignalized crossing. This study collects observational data from more than 3,300 crossing events at 18 intersections in Minnesota to further our understanding of what factors positively influence driver yielding. Using the collected data, a statistical analysis was conducted to identify features that most strongly correlate with driver yielding. Event specific features such as speed were found to greatly influence yielding, with vehicles traveling at a speed of greater than 25 mph significantly less likely to yield to pedestrians than vehicles traveling at speeds lower than 25 mph. Site-specific features such as the presence of signs indicating a crossing were also strongly correlated with driver yielding. The results provide indication of which features of unsignalized crossings correlate with higher driver yielding rates. These findings can be used to guide policy and design at sites where a high driver yielding rate is desirable.					
17. Document Analysis/Descriptors Driving behavior, Yielding, Pedestrians, Unsignalized intersections				18. Availability Statement No restrictions. Document available from: National Technical Information Services, Alexandria, Virginia 22312	
19. Security Class (this report) Unclassified		20. Security Class (this page) Unclassified		21. No. of Pages 91	22. Price

GUIDELINES FOR SAFER PEDESTRIAN CROSSINGS: UNDERSTANDING THE FACTORS THAT POSITIVELY INFLUENCE VEHICLE YIELDING TO PEDESTRIANS AT UNSIGNALIZED INTERSECTIONS

FINAL REPORT

Prepared by:

Raphael Stern
Tianyi Li
Department of Civil, Environmental, and Geo- Engineering
University of Minnesota

June 2023

Published by:

Minnesota Department of Transportation
Office of Research & Innovation
395 John Ireland Boulevard, MS 330
St. Paul, Minnesota 55155-1899

This report represents the results of research conducted by the authors and does not necessarily represent the views or policies of the Minnesota Department of Transportation or the University of Minnesota. This report does not contain a standard or specified technique.

The authors, the Minnesota Department of Transportation, and the University of Minnesota do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to this report.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Technical Liaison Hannah Pritchard for helping to conduct the research and for the many discussions that helped shape this project. Additionally, the authors would like to thank Lisa Bigham (MnDOT), Sara Buermann (Wright County), Bradley Estochen (Ramsey County), Chris Hartzell (City of Woodbury), Bruce Holdhusen (MnDOT), Derek Leuer (MnDOT), Ben Manibog (City of Richfield), James Miles (MnDOT) and Sonja Piper (MnDOT) who served on the Technical Advisory Panel and provided valuable feedback and input during this project. Finally, the authors would like to acknowledge the help of John Hourdos, Melissa Duhn, Joshua Klavins, and John Cullom for their assistance in data collection and data processing.

TABLE OF CONTENTS

CHAPTER 1: Introduction and benefits	1
1.1 introduction	1
1.2 Benefits	1
1.3 Reduced risk.....	2
1.4 Safety	3
1.5 Organization.....	3
CHAPTER 2: Review of relevant literature	4
CHAPTER 3: Site selection	7
3.1 Site characteristics considered	7
3.2 Selected sites	7
CHAPTER 4: Data collection	10
4.1 Video data processing.....	10
4.1.1 Manual data extraction.....	10
4.1.2 Computer-automated data extraction	10
4.2 Collected data	11
CHAPTER 5: Data analysis	14
5.1 Overview	14
5.2 Data format.....	14
5.3 Summary statistics	15
5.3.1 Data Visualization	18
5.4 Analysis Methodology.....	25
5.4.1 Logistic Regression Model	25
5.4.2 Model Evaluation	32
5.4.3 Other models tested	35
CHAPTER 6: Findings	36
6.1 Interpreting model results	36
6.1.1 Interpreting odds ratio.....	36
6.1.2 Model with selected features	36
6.1.3 Model with site features.....	38
6.2 Discussion of factors that impact yielding	38
6.2.1 Significance of signs on driver yielding	39
6.2.2 Significance of adjacent land use on driver yielding.....	40

6.2.3 Significance of adjacent bus stops on driver yielding	41
6.2.4 Significance of the number of lanes on driver yielding	41
6.2.5 Significance of bike lanes on driver yielding	41
6.2.6 Significance of markings on driver yielding	41
6.2.7 Other significant nonsite-specific factors in driver yielding	42
6.2.8 Relationship between crash history and driver yielding rates	44
CHAPTER 7: Recommendations and conclusions	46
7.1 Design recommendations for increasing driver yielding	46
REFERENCES	47
APPENDIX A: Literature review	
APPENDIX B: Data collection	
APPENDIX C: identified features	

LIST OF FIGURES

Figure 1: Map showing the locations of all sites in the Twin Cities.	8
Figure 2: Map showing locations of data collection sites in Northfield, MN.	9
Figure 3. Relationship between number of interactions and driver yielding rates	17
Figure 4. Relationship between vehicle speed and driver yielding rates.	18
Figure 4: Distribution of number of non-yielding vehicles for each pedestrian party at each site.	19
Figure 5: Distribution of pedestrian wait times at each site.....	20
Figure 6: Location vs. yielding outcome showing different yielding rates among sites.	21
Figure 7: Summary of influence of predictor variables on driver yielding response: Pedestrian and vehicle characteristics	22
Figure 8. Summary of influence of predictor variables on driver yielding response: Physical site characteristics	23
Figure 9. Summary of influence of predictor variables on driver yielding response: Signs and markings .	24
Figure 10. Summary of influence of predictor variables on driver yielding response: Adjacent land uses	25
Figure 11: Predictors vs. yielding probability.....	31
Figure 13: ROC curves for different models.....	35

Figure 14: Impact of a sign marking the presence of a pedestrian crossing on the driver yield rate.	40
Figure 15: Impact of number of lanes on driver yield rate.	41
Figure 16: Comparison of driver yield rates for different types of crossing markings.	42
Figure 17: Impact of party size on driver yielding rate. Larger party sizes see a significantly higher driver yield rate.	43
Figure 18: Impact of vehicle type on driver yield rate.	44
Figure 19: Comparison of crashes and driver yield rate showing no significant correlation between the two.	45
Figure 20: TIM unit deployed showing water tight enclosure at bottom containing microcomputer and batteries as well as extendable mast and camera.	1
Figure 21: Data collection site at 22nd Ave NE and University Avenue, Minneapolis.	3
Figure 22: Data collection site at Penn Ave S and W 63rd St, Richfield, MN.	4
Figure 23: Data collection site for Cleveland Ave S and W Pinehurst Ave, St. Paul, MN.	5
Figure 24: Data collection site at 66th and Grand, Richfield, MN.	6
Figure 25: Data collection site at E 66th St. and Elliot Ave, Richfield, MN.	7
Figure 26: Data collection site at E 66th St. and Chicago Ave, Richfield, MN.	8
Figure 27: Data collection site at Portland Ave and E 74th St. Richfield, MN.	9
Figure 28: Data collection site for County B Rd W and Pascal St. Roseville, MN.	10
Figure 29: Data collection site at Division St and 6th St. E, Northfield, MN.	11
Figure 30: Data collection site at Division St and 3rd St. E, Northfield, MN.	12
Figure 31: Data collection site at Nicollet and West 73rd St, Richfield, MN.	13
Figure 32: E 38th St. and 24th Ave, Minneapolis, MN.	14
Figure 33: Data collection site at Snelling Ave S and Stanford Ave, St. Paul, MN.	15
Figure 34: Data collection site at Dale St. and Sherburne Ave, St. Paul, MN.	16
Figure 35: Data collection site at N Upton Ave and Lowry Ave N, Minneapolis, MN.	17
Figure 36: Data collection site at W 52nd St and France Ave S, Edina, MN.	18
Figure 37: Data collection site for 9th Ave NE and Marshall St. NE, Minneapolis, MN.	19

Figure 38: Data collection site at Lowry Ave and James Ave Minneapolis, MN. 20

LIST OF TABLES

Table 1: MnDOT Research Steering Committee Benefits Summary..... 2

Table 2: Factors identified in literature to influence driver yielding behavior. 5

Table 3: Summary of video data collection sites. 12

Table 4: Summary of site yielding statistics..... 16

Table 5: Independent Variables for Logistic Regression Models..... 26

Table 6: Logistic regression model results with selected variables. 29

Table 7: Logistic regression model results with site variables. 32

Table 8: Model performance in test data. 34

Table 9: Odds ratio for different factors indicating change in driver yielding rate for each unit of change in factor..... 37

Table 10: Logistic regression model results with site variables..... 39

EXECUTIVE SUMMARY

Many factors influence an individual driver's decision to yield or not yield to individual pedestrians attempting to cross the road at an unsignalized crossing. These factors include both site-specific factors such as the presence of signs or crosswalk markings as well as event-specific features such as the speed at which the vehicle is traveling. This study collects observational data from more than 3,300 crossing events at 18 intersections in Minnesota to further our understanding of what factors positively influence driver yielding.

Data was collected in the fall of 2020 and the summer of 2021 using custom-built data collection units capable of collecting roughly two weeks of video data from a camera mounted on a 30-foot mast. This allowed for naturalistic data collection and direct observation of driver yielding behavior.

Using the collected data, a statistical analysis was conducted to identify features that most strongly correlate with driver yielding. Event specific features such as speed were found to greatly influence yielding, with vehicles traveling at a speed of greater than 25 mph significantly less likely to yield to pedestrians than vehicles traveling at speeds lower than 25 mph. Site-specific features such as the presence of signs indicating a crossing were also strongly correlated with driver yielding. Interestingly, of the roughly 3,300 pedestrian crossing events, fewer than 4% represented crossings where the driver entered the intersection before a sufficient gap was apparent, suggesting that pedestrians generally wait for a sufficient gap before beginning to cross.

The results provide indication of which features of unsignalized crossings correlate with higher driver yielding rates. These findings can be used to guide policy and design at sites where a high driver yielding rate is desirable.

CHAPTER 1: INTRODUCTION AND BENEFITS

1.1 INTRODUCTION

When a pedestrian tries to cross an unsignalized intersection that a driver is approaching, there are two primary potential outcomes that may occur: the driver can yield, or the pedestrian can yield. Many factors influence the outcome of the interaction between a driver and a pedestrian at an unsignalized intersection. This project aims to gain an understanding of what factors influence the driver yielding rate at unsignalized intersections and hopes this knowledge can be used to inform the design of intersections that encourage drivers to yield to pedestrians.

As part of this research, we collected naturalistic data on driver-pedestrian interactions at 18 intersections in Minnesota. This represented the largest-known publicly available dataset on driver-pedestrian interactions. The video data was manually labeled to identify the outcome of each driver-pedestrian interaction and the data were analyzed using statistical (e.g., logistic regression) analysis to identify the likelihood of a driver yielding to a pedestrian as a function of the variables considered.

With respect to the legal framework around pedestrian safety, Minnesota is considered a *stop for* state. According to Minnesota Statue 169.21, Subdivision 2(a): “where traffic-control signals are not in place or in operation, the driver of a vehicle shall stop to yield the right-of-way to a pedestrian crossing the roadway within a marked crosswalk or at an intersection with no marked crosswalk.” Thus, failure to yield to a pedestrian at an unsignalized intersection is considered a moving violation in Minnesota.

1.2 BENEFITS

Based on the MnDOT Research Steering Committee Criteria in Table 1, this project has two primary benefits: reduced risk and increased safety. Both benefits can be quantified by measuring the reduction in expected non-yielding events at a particular intersection in a set of intersections after identifying what features of the intersection most significantly contribute to the driver yielding rate.

Table 1: MnDOT Research Steering Committee Benefits Summary.

Benefit category	Applicable	Can quantify	How benefits are quantified
Construction savings			
Decreased engineering cost			
Environmental aspects			
Improved lifecycle costs			
Operation/maintenance saving			
Reduce risk	X	Yes	Measure reduction in expected non-yielding
Reduce road user cost			
Safety	X	Yes	Measure reduction in expected non-yielding
Technology			

1.3 REDUCED RISK

This project provides benefit by identifying how to reduce risk to pedestrians in Minnesota. By collecting data on how driver-pedestrians interactions occur, we can identify the key factors in terms of intersection layout, land use, and signs, among other features, that are most strongly correlate with driver yielding. Identifying these features will allow for retrofitting existing intersections and designing new intersections that maximize the likelihood that the driver will yield to a pedestrian.

The reduction in risk for pedestrians comes from pedestrians navigating intersections that are designed to maximize the probability of drivers yielding. Using the experimental data collected, a data-driven model will be constructed that predicts the probability of drivers yielding at a particular intersection based on the underlying features of the intersection.

1.4 SAFETY

By providing tools to engineers at both the state and local level to help design and retrofit intersections so as to reduce risk to pedestrians, this project facilitates pedestrian safety. This can be quantified as the increase in driver yielding that would be expected if any treatments are selected at a particular intersection.

1.5 ORGANIZATION

The remainder of this report is organized as follows: Relevant literature is reviewed in Chapter 2. The criteria for study site selection, and the specific sites identified are discussed in Chapter 3. The data collection process is presented in Chapter 4, and an analysis of the data is conducted in Chapter 5. Findings are presented in Chapter 6, and recommendations for improvements to intersections to increase driver yielding are presented in Chapter 7. Appendix A provides more details on the literature review, Appendix B provides specifics on each individual site selected, and Appendix C presents the important intersection and event features used for analysis.

CHAPTER 2: REVIEW OF RELEVANT LITERATURE

The National Highway Traffic Safety Administration [1] recorded a total of 6,516 pedestrian fatalities in 2020, a 5% increase over 2019 and an 8% increase over 2017. This is the highest number of pedestrian fatalities since 1989. A total of 38,824 people died in traffic crashes in 2020, and the risk of injury and death disproportionately falls on pedestrians. For example, in 2017, less than 1% of miles traveled were on foot, yet 16% of traffic fatalities were pedestrians.

This risk to pedestrians is especially elevated when crossing the road, where 91% of pedestrian fatalities occur. Understanding the factors that lead to higher rates of pedestrian-involved crashes has been of interest for a long time with early works dating back to the 1970s. For example, Herms [2] studied crashes that happened at 400 intersections over a 5-year period, each intersection with one marked and one unmarked crosswalk crossing the main thoroughfare to compare the driver-pedestrian interactions at these crossings. The author found 177 people were hit in the 400 marked crosswalks while 31 pedestrians were hit in the 400 unmarked crosswalks, 18 fatalities occurred in the marked crosswalks, and 3 fatalities in the unmarked crosswalks. Without accounting for many externalities, and with no information about the pedestrians' perceptions, the author ascribed this finding to the pedestrians' perception of safety when using the marked crosswalk. However, this may not account for the fact that generally speaking, crosswalks typically are located in places with high pedestrian volume, and thus may be locations where a crash is more likely.

As shown by the previous example, these early efforts as well as coverage of pedestrian fatalities in news media often focused on pedestrian fault in how the interaction and resulting crash or crash risk was described [3]. The authors analyzed the text from 200 local news articles to answer how articles apportion blame between pedestrians and drivers. They find that blame is often placed on the pedestrian, even when data on a pedestrian's motivations or perceptions is not available. This bias is consistent with what is found in many research studies.

A comprehensive literature review is included in Appendix A. A summary of the relevant literature is provided in Table 2 below. This table provides a summary of important features for driver yielding that have been identified across the literature. These previous findings influenced the experiment design and data analysis in this study.

Table 2: Factors identified in literature to influence driver yielding behavior.

Event Features		Site Features	
Feature Name	Literatures	Feature Name	Literatures
Time	[4], [5]	Posted Speed	[6]
Party Size	[7], [8], [9]	Number of Lanes	[7], [10]
Pedestrian Position	[7], [11]	Signage	[8], [9], [12]
Pedestrian Type	[10], [13], [14]	Markings	[4]
Vehicle Speed	[6], [9], [10], [11]	Land Use	[15]
Opposite Direction	[8], [9]	Lightning	[4], [16]
Following Vehicles	[9], [4]	Road Surface	[6], [4]
Vehicle Type	[9], [4]	Bus Stops	[6]
Type of Yield	[8]	Traffic Volume	[6], [15]
Weather	[4]	Bike Facility	[17], [18], [19]
Vehicle Turning Direction	[20]	Intersection Angle	[21]
Pedestrian gender	[22], [13]	Lane width	[23]

In summary, most of the studies mentioned the importance of vehicle speed, which is often seen as the most direct influence on the outcome of a pedestrian-vehicle interaction. Experimentally, many studies used radar to record vehicle speed, while others rely on computer vision technologies to extract vehicle speed.

The driver's yielding decision is also influenced by the presence of other vehicles, such as following vehicle and whether the opposite direction vehicle is yielding, the driving direction, and the type of vehicle also have effects. From the pedestrian's perspective, the pedestrian volume, party size, gender, and type of pedestrian are factors that are identified in the literature as important. The time and weather of the event happening are also documented in some studies.

Another area of focus is site-specific features. As shown in Table 2, the signs and markings of the intersection are important variables for the vehicle yielding decision. The built environment, land use, and nearby bus stops are also mentioned as important features, along with demographic features. Traffic volume (e.g., AADT) is identified as one of the most important factors for traffic safety as well [38].

At the macroscopic level, many studies [20, 21, 23, 25] argued that the more pedestrians are present on the road, the safer they will be. Thus, the challenge is reaching the critical mass of pedestrians, where they will feel safe, since beyond that, the presence of pedestrians will likely draw additional pedestrians on the road. Thus, efforts made to improve pedestrian safety will likely have additional benefits, since safer roads will likely lead to the presence of more pedestrians, which, in turn, will continue to improve safety for pedestrians.

CHAPTER 3: SITE SELECTION

This chapter presents the rationale behind site selection for this study. These sites were selected based on conversations with the project Technical Advisory Panel (TAP). Many candidate sites were identified, but several of these sites were not suitable for the study based on specific site restrictions (e.g., ongoing construction or confounding factors such as poor visibility at the site). The resulting final list of sites for data collection was approved by the TAP before data collection began.

3.1 SITE CHARACTERISTICS CONSIDERED

Sites were selected to represent a variety of conditions along five primary dimensions: posted speed limit, number of lanes, presence of signs at the crossing, presence (and type) of markings, and land use context. Within each of these categories, the TAP expressed interest in identifying sites that had variability in each dimension, as well as a portfolio of sites that the group suspected would exhibit a diversity of driver yielding rates. Thus, sites were selected that, based on anecdotal experience, had both low and high driver yielding rates. The motivation being that a variety of yielding rates would help provide clues to which intersection features result in higher driver yielding rates.

Certain site characteristics such as sites with RRFBs were specifically excluded due to the extensive research on their impacts on driver yielding behavior. Additionally, sites with one-way streets were also excluded, since very few one-way streets exist in the MnDOT road network.

3.2 SELECTED SITES

The selected sites are detailed in Table 3. The Twin Cities sites are shown on a map in Figure 1, while the two sites in Northfield, MN are shown on a map in Figure 2. These sites represent a diverse set of sites that were selected based on both the characteristics of the site and practical considerations such as the proximity of adequate poles to mount the data collection units.

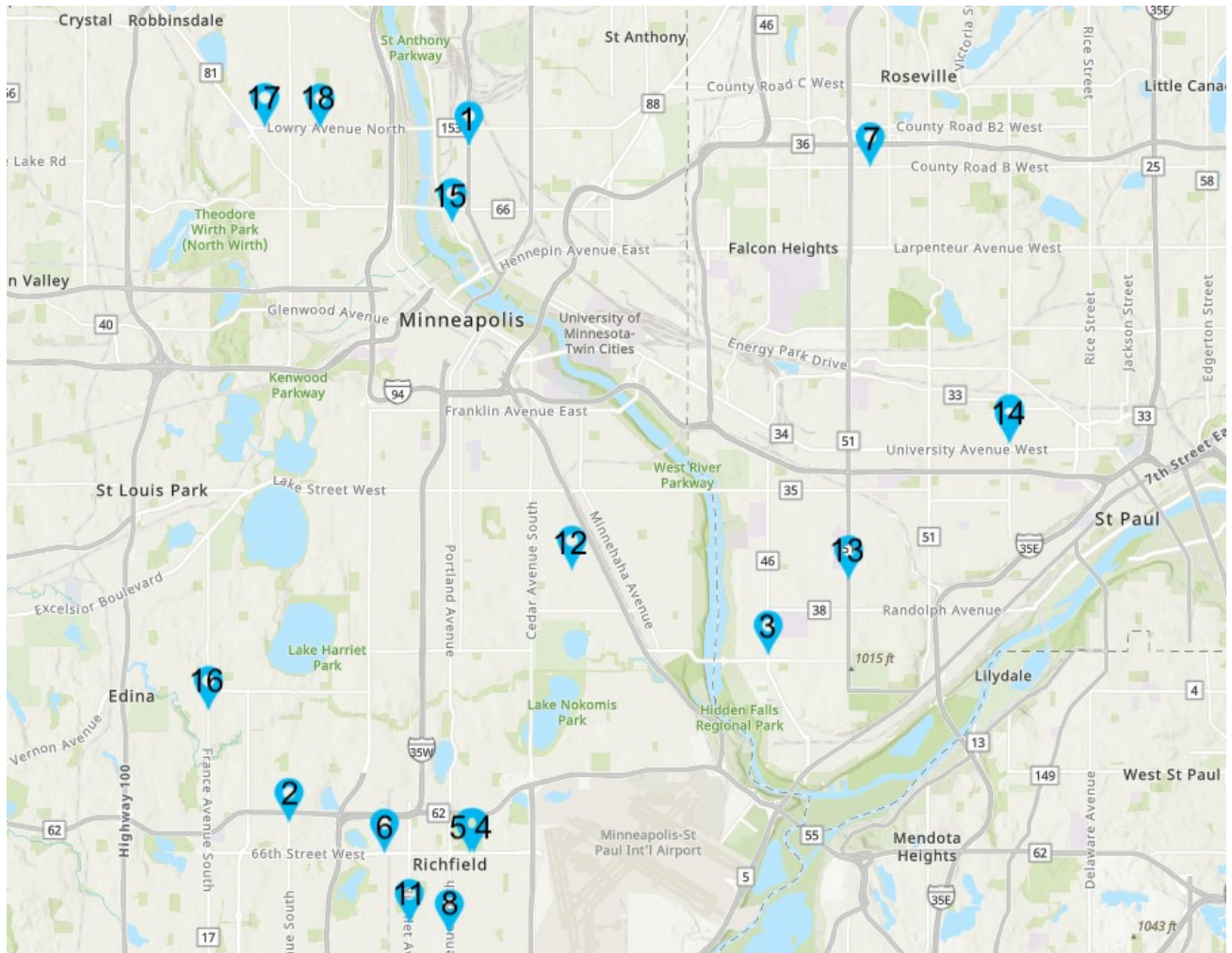


Figure 1: Map showing the locations of all sites in the Twin Cities.

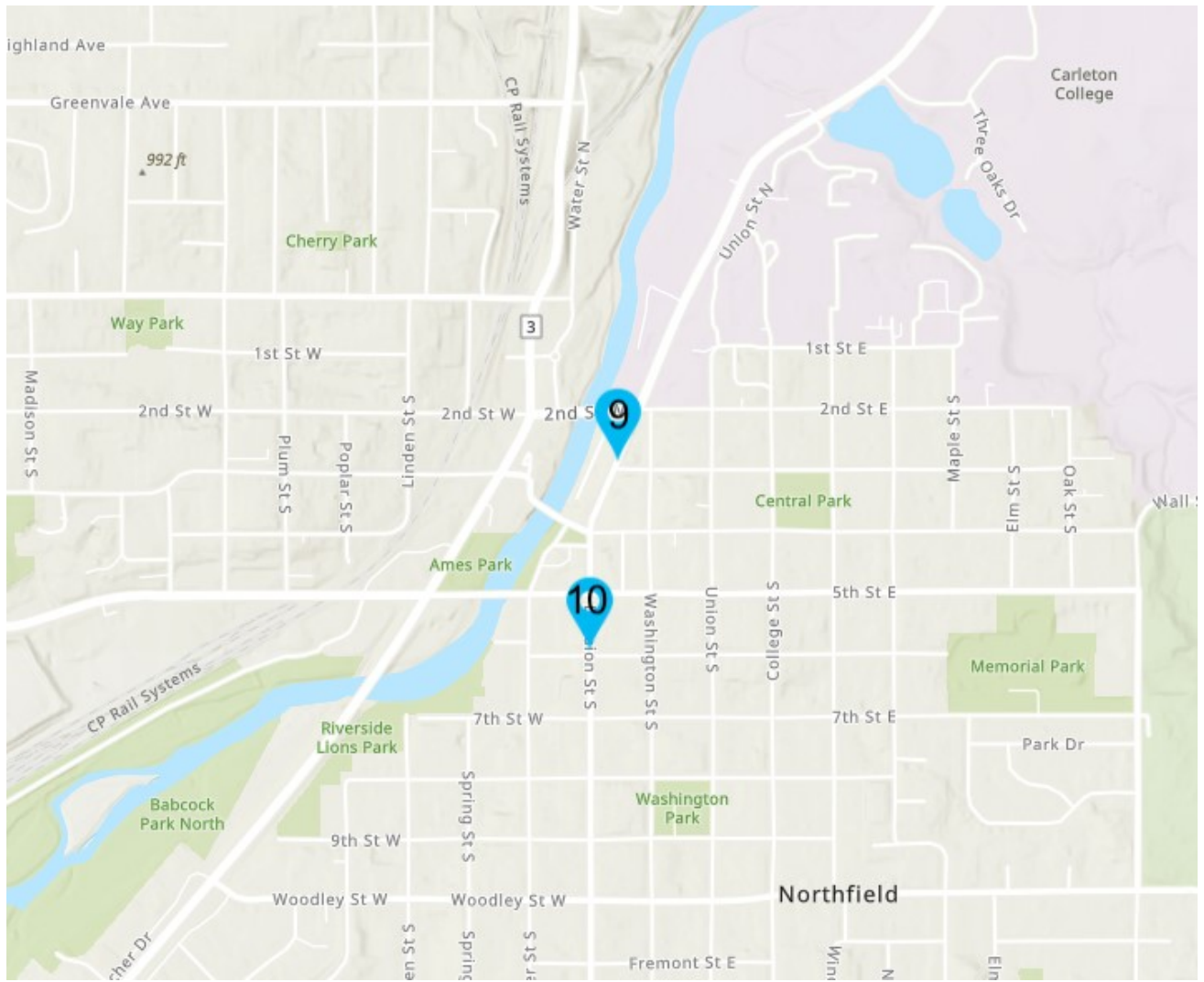


Figure 2: Map showing locations of data collection sites in Northfield, MN.

CHAPTER 4: DATA COLLECTION

Video data were collected using the Minnesota Traffic Observatory (MTO) Traffic Information Monitors (TIMs). TIMs were deployed for two weeks at a time to maximize the potential number of pedestrian-driver interactions observed at each deployment site. The data that were collected at each intersection included video data as well as observational notes about the intersection, TIM location, and relevant intersection characteristics. The collected video data were reduced to numeric data using both video processing techniques as well as manual data extraction. See Appendix B for detailed descriptions of the TIMs, TIMs deployment, and data extraction/reduction.

4.1 VIDEO DATA PROCESSING

The collected video data were reduced to numeric data using both video processing techniques as well as manual data extraction. Both techniques are described below.

4.1.1 Manual data extraction

Manual data collection was conducted by University researchers who viewed the video and identified the clock time of each pedestrian crossing event. The researchers then analyzed each individual crossing event to identify the number of vehicles until a driver yielded as well as the time of day and duration of the crossing along with other crossing characteristics such as party size, type of pedestrian (e.g., pedestrian with child, pedestrian with dog, etc.). Additionally, other interaction specific features such as the vehicle type and the state of the oncoming traffic were identified. See Section 5.2 for discussion of some of these variables that required judgment on the part of the research team.

Quality control for the extracted data was conducted by having additional team members view portions of the video data and extract interactions. These data were used to compare to the original dataset and ensure that the research team was consistent about data extraction.

4.1.2 Computer-automated data extraction

Computer vision is used to extract vehicle speed from the video data. This is done by identifying two virtual trip lines a known distance apart in the video frame. Image detection is used to identify a bounding box around each vehicle, and the time difference between crossing the first and second trip line is used to estimate the travel time. This is then used to estimate the average approach speed and added to the comprehensive data spreadsheet.

As with any data extraction method, the computer-automated methods are susceptible to some types of faults. For example, there are some circumstances under which the computer is unable to or incorrectly identifies the trajectory of the vehicle. This may be the result of passing vehicles, occlusion, or other lighting conditions such as shadows that make it difficult to properly identify the outline of the vehicle. Quality control was conducted by manually inspecting the identified vehicle trajectories to ensure they are accurate.

4.2 COLLECTED DATA

The data collection took place at 18 intersections in Minnesota. Data collection occurred in two phases: Phase I was a pilot data collection with 3 sites that took place in the fall of 2020, after which the data collection procedure was refined. Phase II took place in the summer of 2021, and included a total of 15 sites.

Aerial maps of each of the data collection sites and notes describing the TIMs placement and any other site specific issues are provided in Appendix B.

Table 3: Summary of video data collection sites.

Site Number	City	Intersection	Data collection dates
1	Minneapolis	22 nd Ave NE and University Ave	9/21/2020 – 10/2/2020
2	Richfield	Penn Ave S and W 63 rd St	9/21/2020 – 10/2/2020
3	St. Paul	Cleveland Ave and Pinehurst Ave	9/23/2020 – 10/2/2020
4	Richfield	66 th and Elliot	5/28/2021 – 6/15/2021
5	Richfield	66 th and Chicago	5/28/2021 – 6/15/2021
6	Richfield	66 th and Grand	5/28/2021 – 6/15/2021
7	Roseville	County B Rd W and Pascal St	6/3/2021 – 6/11/2021
8	Richfield	Portland Ave S and 74 th St	6/1/2021 – 6/15/2021
9	Northfield	3 rd and Division	6/15/2021 – 6/30/2021
10	Northfield	6 th St E and Division Ave	6/15/2021 – 6/30/2021
11	Richfield	Nicollet and W 73 rd	6/16/2021 – 6/30/2021
12	Minneapolis	E 38 th St and S 24 th Ave	6/30/2021 – 7/11/2021
13	St. Paul	Snelling Ave S and Stanford Ave	7/1/2021 – 7/15/2021
14	St. Paul	Dale St N and Sherburne Ave	7/1/2021 – 7/15/2021
15	Minneapolis	9 th Ave NE and Marshall St NE	7/15/2021 – 7/29/2021

16	Edina	W 52 nd St and France Ave S	7/15/2021 – 7/29/2021
17	Minneapolis	N Upton Ave and Lowry Ave N	7/15/2021 – 7/29/2021
18	Minneapolis	Lowery Ave N and N James Ave	9/25/2021 – 10/5/2021

CHAPTER 5: DATA ANALYSIS

5.1 OVERVIEW

This chapter explains the extracted data including the data format and individual characteristics recorded, along with a detailed description of each data feature. Additionally, we provide a summary of the data collected and a detailed analysis of the data including the construction of logistic regression and other machine learning models to predict driver yielding behavior based on the observed data.

5.2 DATA FORMAT

This section describes how the data extracted from the videos is formatted. All data is provided in the `ped_yielding_data.xls` file which is publicly available for download: https://z.umn.edu/pedestrian_data. Each pedestrian crossing event is listed as a new row. The entry in each column represents a different feature of the interaction as described below.

The features identified for each driver-pedestrian interaction are provided were generally grouped into two broad categories: features that are specific to the individual interaction and features that are specific to the individual site. These features include observations about the individual interaction such as the time of day, the number of pedestrians and the type of vehicle involved in the interaction as well as site-specific features such as the number of lanes, the crossing width, and the presence of marked crosswalks or signs. The full list of features identified is presented in Appendix C. Individual variables that required the judgement of the research team are discussed in more detail below:

- **Close Call:** This feature was identified at the discretion of the research team and was used to identify potentially dangerous crossings, based on the research team's judgement. This feature was mostly identified to make future analysis on close calls and near miss crossings more straightforward using the collected data. The data is coded as 1 if the crossing is perceived as a close call and 0 if not. There were very few close calls at any of the 18 sites.
- **Interaction/Event Type:** This feature was used to identify the nature of the interaction. The coding was based on the judgement of the research team and was decided based on discussions amongst the researchers looking at individual crossings. Note that when The coding was as follows: A: pedestrian(s) crossed at a comfortable pace as soon as a gap was present; B: pedestrian(s) crossed at fast pace to utilize small gap as soon as a gap was present; C: pedestrian(s) waited after gap was present before crossing at a comfortable pace; D: pedestrian(s) waited after gap was present before crossing at fast pace; E: pedestrian(s) did not wait for gap before crossing.
- **Type of Yield:** S for a vehicle that comes to a complete stop to allow the pedestrian(s) to cross, R for a rolling yield where the driver decreases speed but does not come to a complete stop, O for any other form of yielding (e.g., changing driving path to avoid pedestrian(s)), N or blank entry for a non-yielding driver.

- **Comments:** This field was included to allow for any additional comments as noted during data extraction. This field was mostly used to call attention to individual unique driver-pedestrian interactions that may require further analysis in future studies.

The distribution between categories for the interaction/event type and type of yield features ended up telling an interesting story all on their own. Many practitioners work under the assumption that pedestrians who are involved in a crash simply did not wait for an appropriate gap before crossing. This belief may be based on an assertion about “false sense of security” from one of the first crosswalk studies [2]. The idea is also reinforced by survivor bias in crash reporting, as drivers often report that a pedestrian “came out of nowhere.” However, pedestrian behavior observed in this study contradicts this common assumption. Specifically, of the 3,314 total interactions, only 135 pedestrian began crossing before a gap was fully available (event type E representing only 4% of the cases) and 764 pedestrians waited even after a gap was present before beginning to cross (event type C representing over 34%).

5.3 SUMMARY STATISTICS

Table 1 presents the summary statistics of the 18 intersections studied. There were 3,314 total interactions between a driver and one or more pedestrians across the 18 intersections that made up the data set for the analysis.

- Intersection 16 had the most interactions with 840 events, and intersection 11 had the least interactions with only 20.
- Intersection 16 also had the highest yielding rate of 70.36%.
- Intersections 11 and 15 had the lowest yielding rate of 0%, meaning we did not observe a single driver yield to a pedestrian at these locations.

Table 4: Summary of site yielding statistics

SITE	TOTAL INTERACTIONS	DRIVER YIELD RATE (%)	AVERAGE PEDESTRIAN WAIT TIME
1	31	16.13%	4.65
2	100	43%	9.66
3	640	67.87%	4.97
4	143	5.59%	10.57
5	341	2.93%	15.81
6	95	17.89%	9.98
7	121	14.50%	14.15
8	105	16.19%	10.27
9	74	5.41%	11.39
10	88	11.36%	10.17
11	20	0%	13.7
12	122	4.92%	12.58
13	98	59.18%	5.29
14	133	64.47%	5.04
15	46	0%	18.52
16	840	70.36%	7.74

17	127	17.32%	11.84
18	189	5.29%	7.05
ALL	3314	40.38%	9.43

Overall, the data show that the number of pedestrians crossing at intersections is found to correlate with higher driver yielding rates. Note that the wait time listed in Table 4 include the wait time for all pedestrians. The average wait time for pedestrians who had to wait for a vehicle to yield was only slightly higher across all sites with an average of 10.2 s. Intersections 3 and 16 had the highest number of interactions and the highest driver yielding rate. This is intuitive, since drivers who use the intersection frequently may be more likely to expect pedestrians and more prepared to yield.

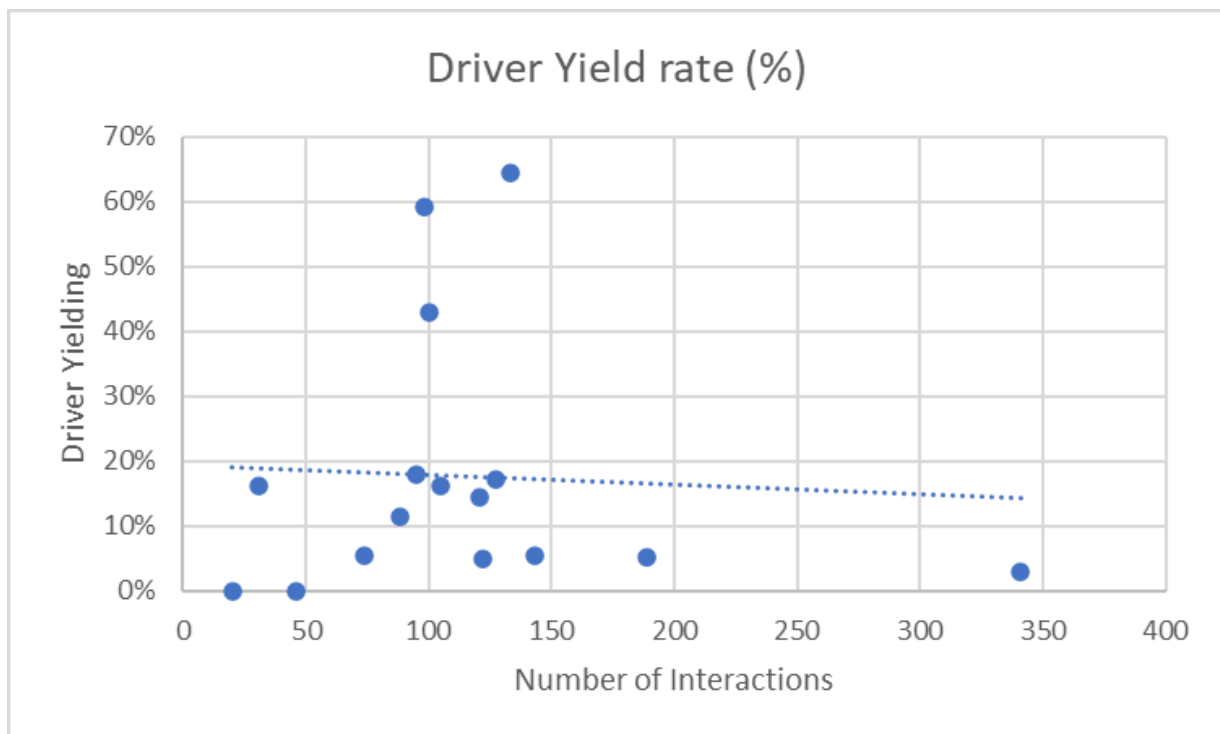


Figure 3. Relationship between number of interactions and driver yielding rates removing two outlier sites with high pedestrian volumes. No clear correlation between number of interactions and driver yield rate is observed.

Information about driver speed was also collected at all 18 sites. As expected, as vehicle speed increased, the percentage of interactions in which drivers yielded to pedestrians decreased. This nearly linear relationship is clearly seen in Figure 4.

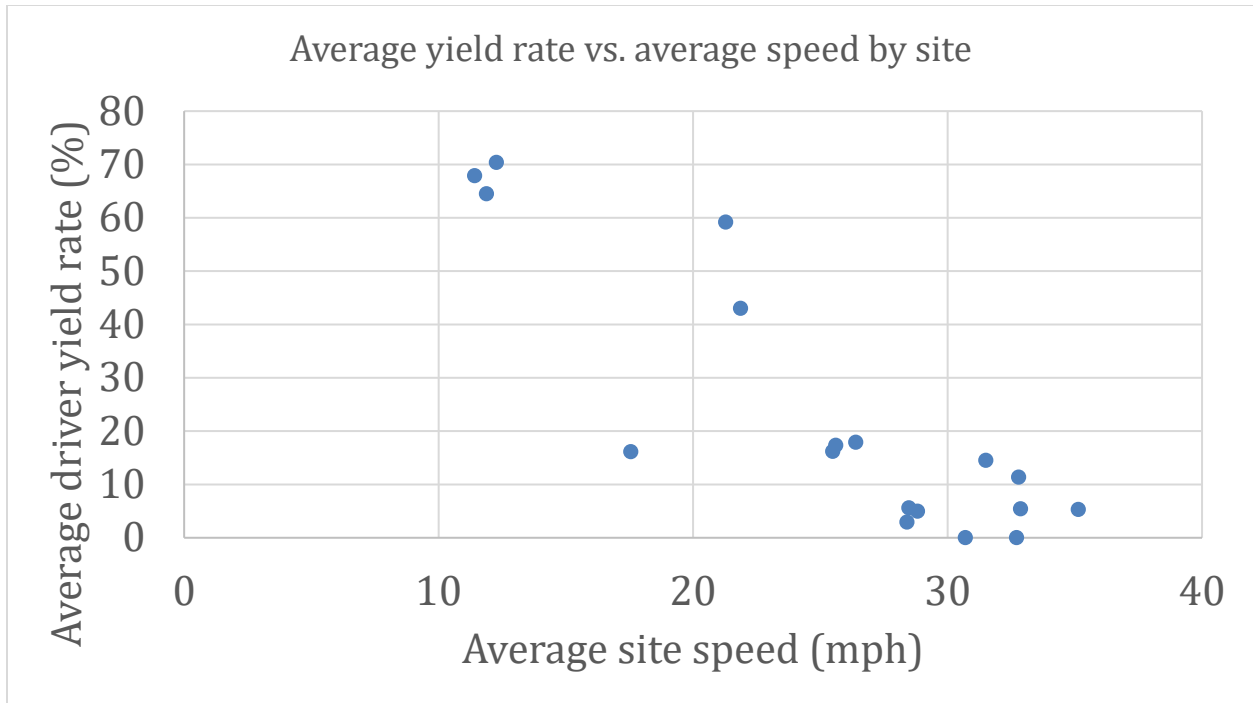


Figure 4. Relationship between vehicle speed and driver yielding rates.

5.3.1 Data Visualization

In total, the extracted data consists of 3,314 interactions from the 18 intersections. No collisions are observed in the dataset, although there were 16 close calls. The vehicle yielding rate observed varies among the sites as seen in Figure 5 and the distribution of wait times for pedestrians to cross for different sites is presented in Figure 6. In both figures the site number corresponds to the site number in 0.

The distribution of non-yielding vehicles at each site shown in Figure 5 shows that across sites, generally pedestrians are able to cross relatively quickly. This is also seen in the distribution of wait times in Figure 6. However, the variability in both waiting time and number of non-yielding vehicles across sites suggests that individual sites have fundamentally different driver yielding characteristics.

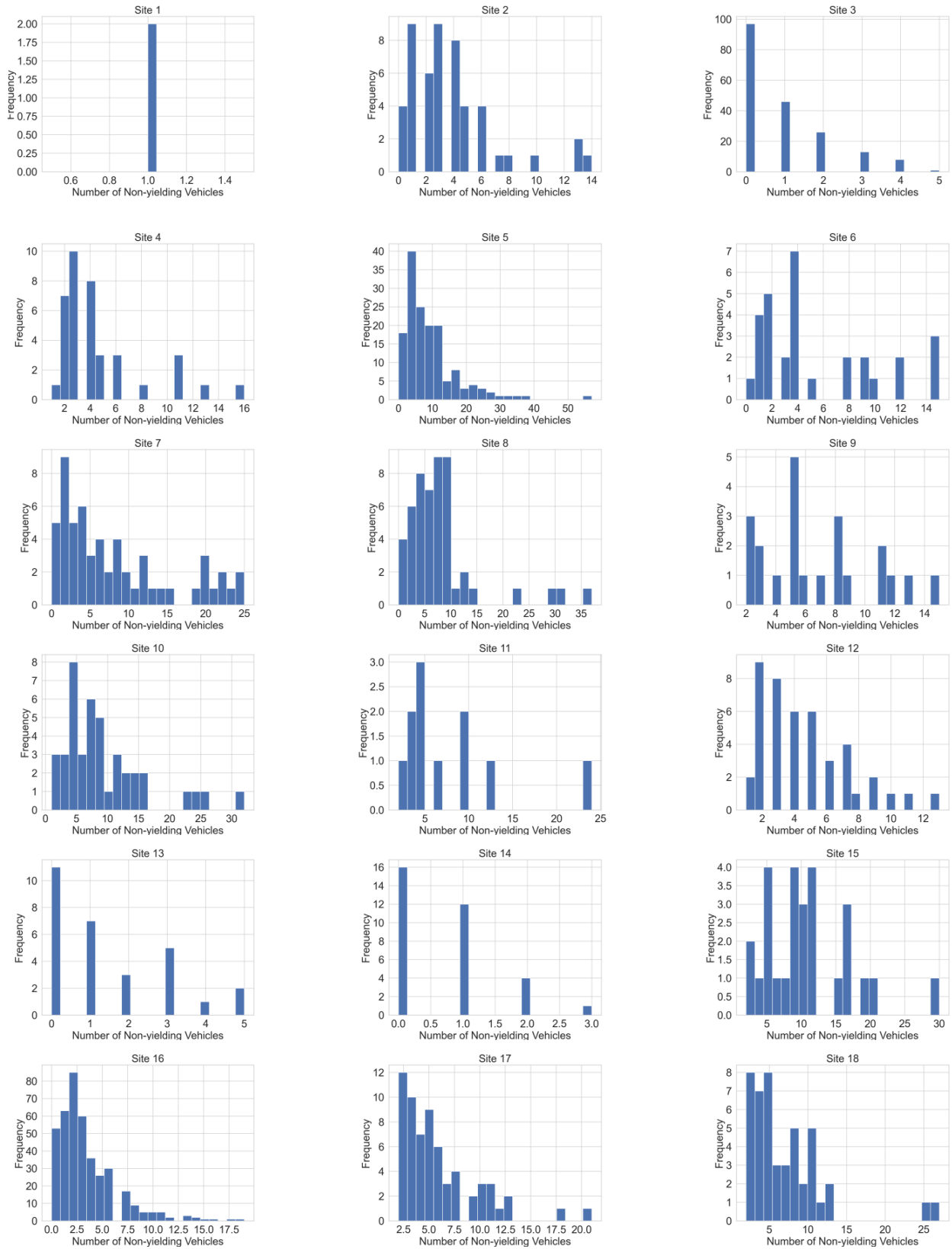


Figure 5: Distribution of number of non-yielding vehicles for each pedestrian party at each site.

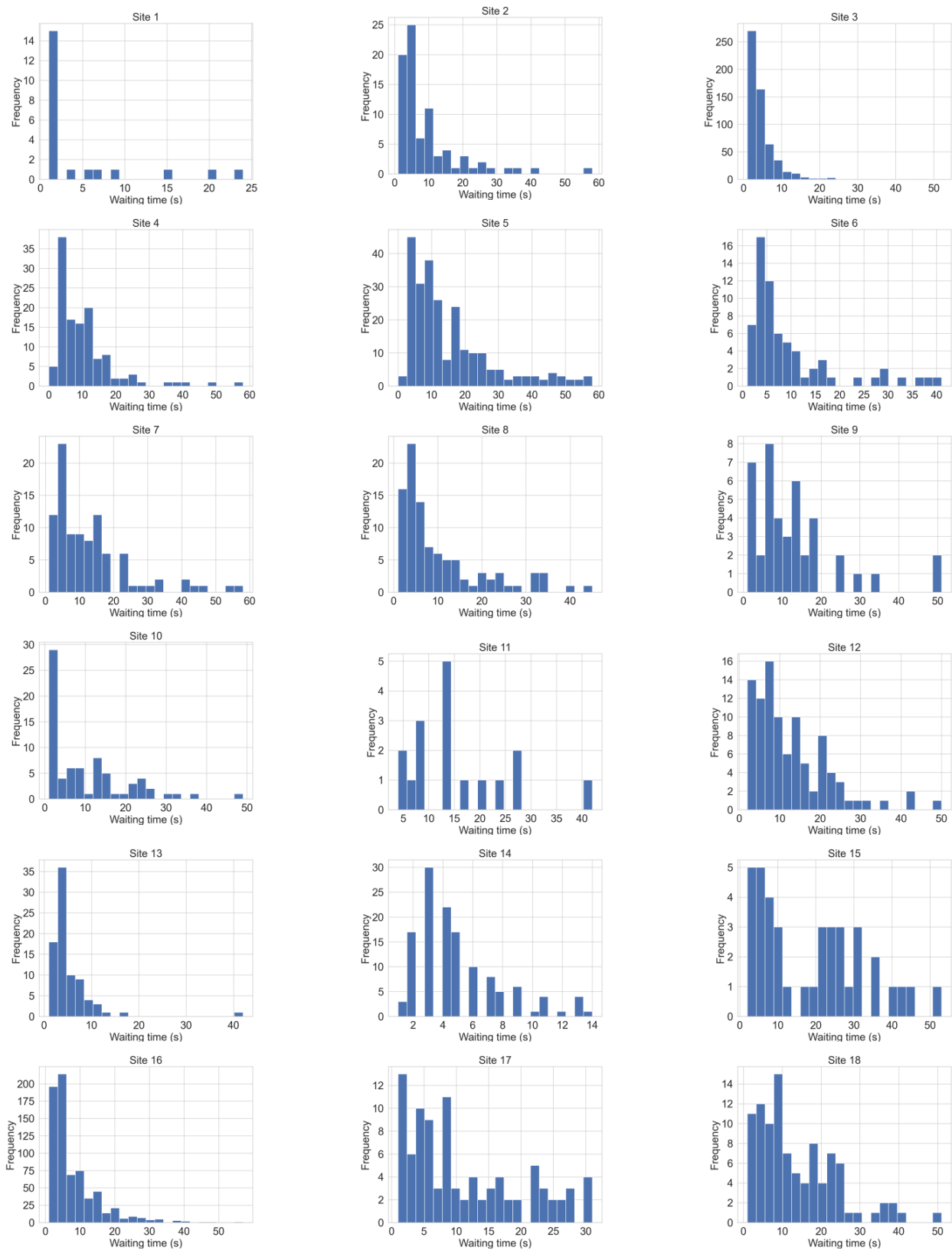


Figure 6: Distribution of pedestrian wait times at each site.

As illustrated in Figure 7, the yielding behavior varies a lot among locations. Note that the location number corresponds to the location number identified in Table 3. As discussed, Sites 3 and 16 both had a larger total number of interactions than any other site as well as a higher ratio of drivers yielding compared to pedestrians yielding. Meanwhile, sites 4, 5 are a block apart on the same corridor (66th Street, Richfield). Both sites are three-way intersections along a park, but only site 5 has a bus stop. We saw more total interactions at site 5 (the third highest volume of the 18 sites), which is intuitive. However, in this case the ratio between drivers yielding and pedestrians yielding was very low at both sites. Site 5 appears to break from the trend of more interactions resulting in better driver yielding.

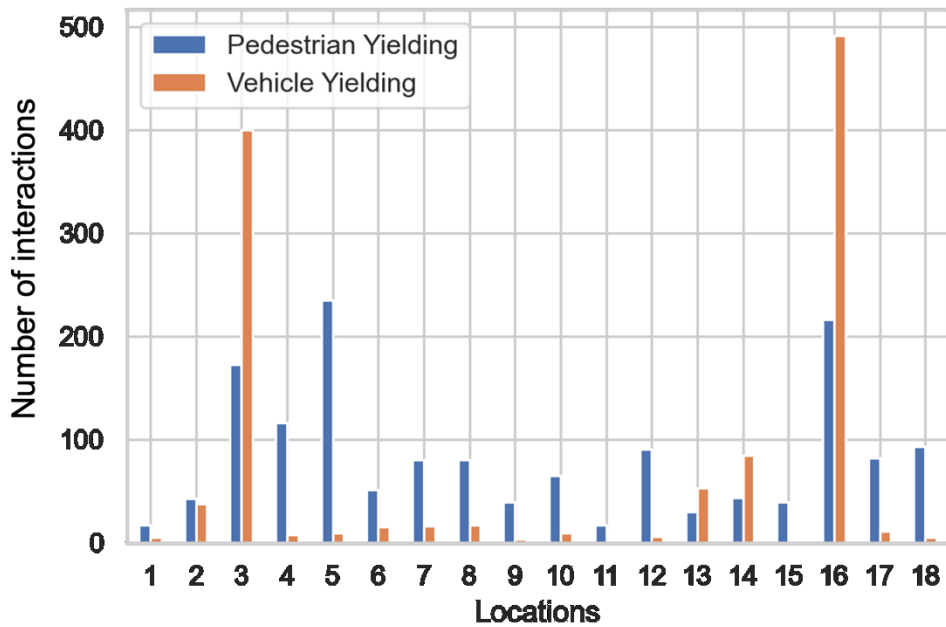


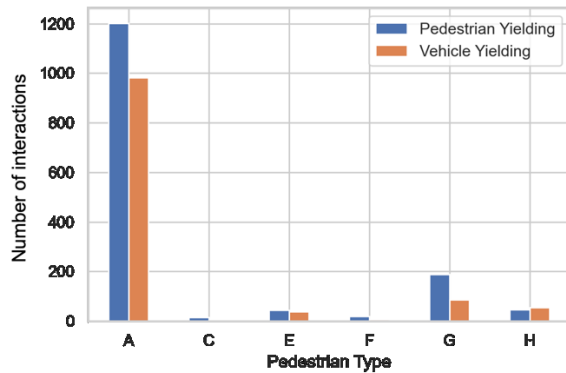
Figure 7: Location vs. yielding outcome showing different yielding rates among sites.

The following figures provide a summary of the collected yielding data and the respective yielding vs. non-yielding numbers for different factors that were found to be particularly informative in understanding the data. While there are clear correlations between these factors, this does not imply causation. I.e., simply because two quantities are correlated does not imply that the presence of one implies the presence of the other.

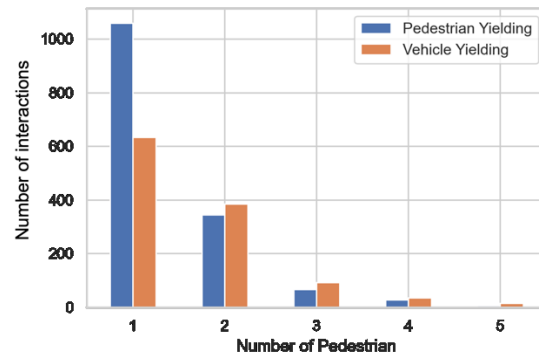
Figure 8 describes the influence of the type of pedestrian, group size, and vehicle type on yielding.

- We classified people crossing into one of eight different categories: A: *person on foot*; B: *person on bike*; C: *person on mobility device (e.g., scooter/hoverboard)*; D: *person walking bike*; E: *mix of pedestrian types*; F: *other*; G: *person with a dog*; H: *person with a stroller or small child*. Note that class B and D are included in the overall data but omitted from this analysis and Figure 8 for since only pedestrian data is of interest for this analysis, but the remaining data is collected and kept for consistency. The most common type was pedestrian on foot and the second most common was pedestrian with a dog.

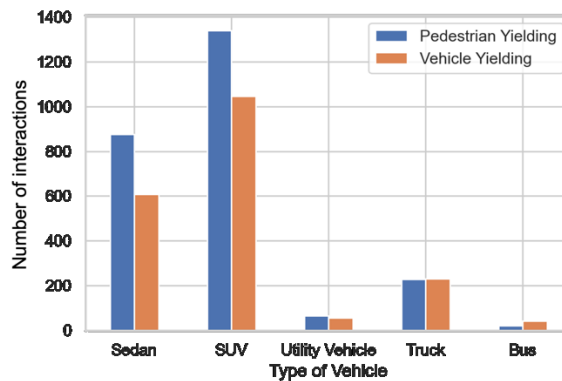
- Yielding behavior of drivers is influenced by party size. The results indicate that drivers are more likely to yield a group of pedestrians than to an individual pedestrian trying to cross the street.
- Drivers of SUVs and sedans are less likely to yield compared with other types of vehicles. Buses and trucks have the highest yielding rate.



(a) Type of pedestrian vs. yielding.



(b) Number of pedestrian vs. yielding.



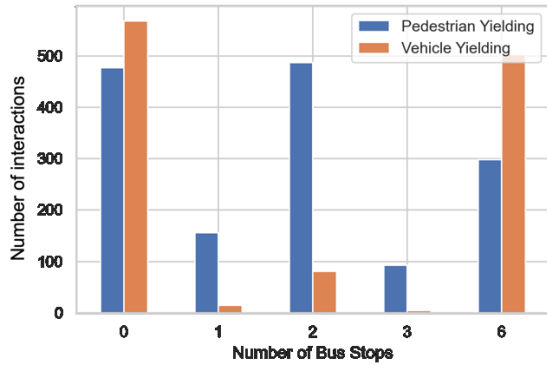
(c) Type of vehicles vs. yielding.

Figure 8: Summary of influence of predictor variables on driver yielding response: Pedestrian and vehicle characteristics

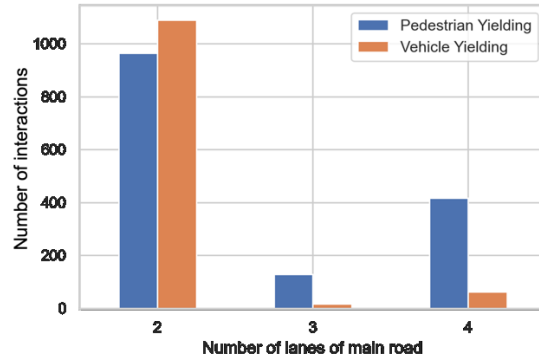
Figure 9 describes the influence of the various physical site characteristics (bus stops, number of lanes, and trees) on yielding.

- If there is at least one bus stop, the number of bus stops is correlated with yielding rates. We find that the vehicle yielding rate is significantly higher when there are more bus stops near the intersection.
- The results indicate that drivers are less likely to yield when there are more lanes on the road.
- Lower driver yielding rates are correlated with the presence of trees obstructing the drivers' view of a waiting pedestrian. Specifically, the results indicate that the highest driver yielding rates are observed at intersections without any trees obstructing the crosswalk, while the

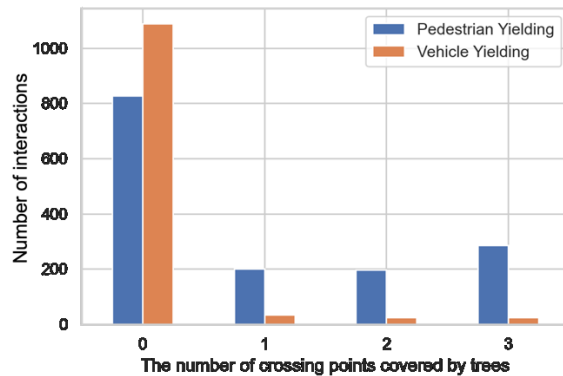
lowest yielding rate is observed for intersections where all four intersection corners have tree cover.



(a) Number of bus stops vs. yielding.



(b) Number of lanes vs. yielding.

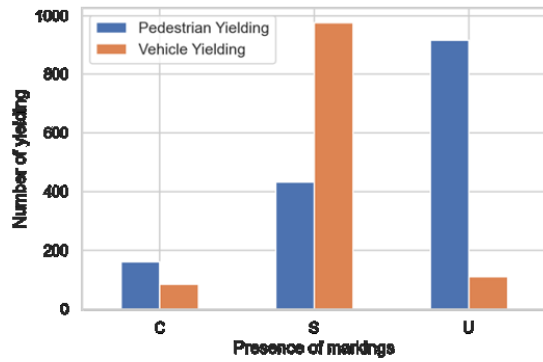


(c) Tree cover vs. yielding.

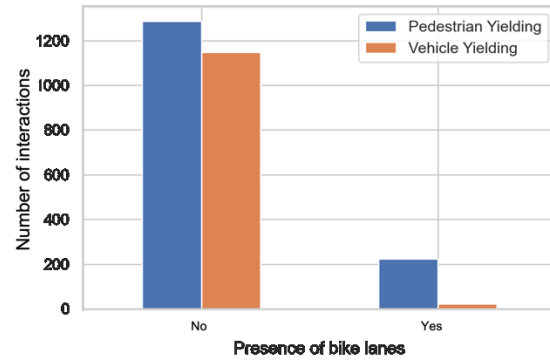
Figure 9. Summary of influence of predictor variables on driver yielding response: Physical site characteristics

Figure 10 describes the influence of the several traffic control features (crosswalk markings, crosswalk signs, and bike lane markings) on yielding.

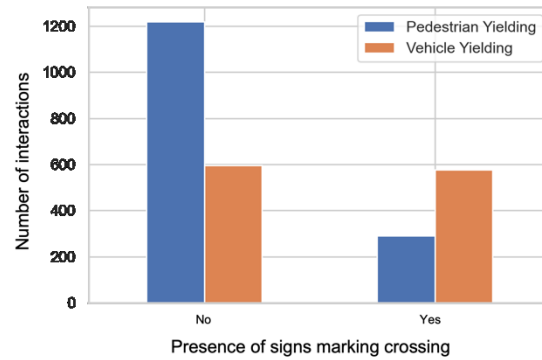
- Sites with crosswalk markings had higher driver yielding rates, which was expected. However, sites with standard markings (two parallel lines) had higher yielding rates than sites with continental markings, which was counter to our expectations. There were 12 sites with no markings, 3 sites with standard markings and 3 sites with continental markings. Note the small sample size of sites with standard and continental markings.
- Sites with crosswalk signs had higher driver yielding rates, which was expected.
- The presence of a bike lane was correlated with a decrease in driver yielding, although only 9.1% of the interactions in the data set occurred at sites with bike lanes.



(a) Type of markings vs yielding.



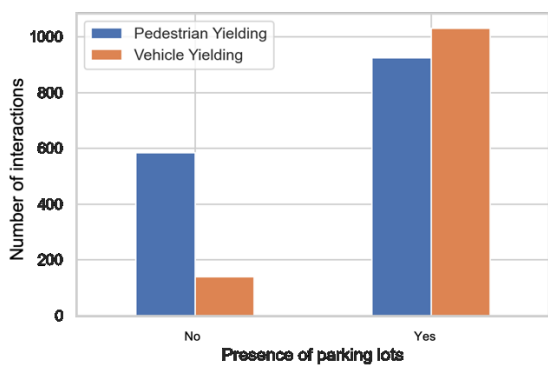
(b) Presence of bike lanes vs. yielding.



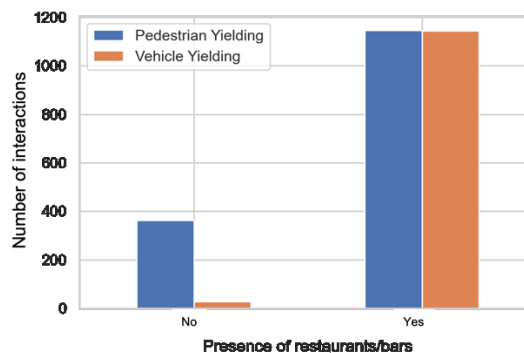
(c) Presence of crosswalk signs vs. yielding.

Figure 10. Summary of influence of predictor variables on driver yielding response: Signs and markings

Finally, several adjacent land use types were recorded during data extraction, including distance to a park or school, presence of multi-family housing. Figure 11 shows the influence of two of these land uses – restaurants/bars and parking lots. Sites with restaurants/bars and/or parking lots represented a large portion of the data set and those land uses were correlated with a higher rate of driver yielding (although in the case of parking lots, pedestrians still yielded more often).



(a) Presence of parking lots vs. yielding.



(b) Presence of restaurants/bars vs. yielding.

Figure 11. Summary of influence of predictor variables on driver yielding response: Adjacent land uses

Importantly, this summary analysis can only identify correlations between individual variables and the outcome of the driver/pedestrian interaction. The analysis that follows will be used to identify how significant this correlation is. In each case, statistical and regression analysis will be used to identify which factors are found to be most important in determining driver yielding and how those factors interact. However, correlation does not imply causation, and many correlations may be identified due to other confounding factors. Therefore, only strong correlations should be considered meaningful, and care must be taken not to place too much weight on correlated quantities.

5.4 ANALYSIS METHODOLOGY

In this section, we present the statistical analysis to model driver yielding rates at the 18 unsignalized intersections studied. Using the developed models, we are able to identify significant factors in driver yielding at these sites and describe the likelihood of a driver yielding given the prevailing conditions at similar sites.

5.4.1 Logistic Regression Model

Logistic regression has been shown to have robust performance in binary classification and provides a model with good interpretability [24]. Logistic regression can be used to model the probability of a certain event of a binary dependent variable Y . In this case, the driver yielding response has a binary outcome where the driver either yields to the pedestrian or fails to yield¹. We assume a linear relationship between the independent variable and the log-odds of the event. This linear relationship is shown in the Equation (1), where L is the log-odds, β_i are the model parameters, and $P(Y = 1)$ for driver

¹ Several types of yielding were considered as outlined in 8.3 Appendix C. However, all were considered as yielding events.

yielding and $1-P(Y = 1)$ for driver not yielding are the binary outcomes. The logistic regression model takes the form in Equation (2) for the probability of driver yielding $P(Y = 1)$, and conversely the probability of the outcome of pedestrian yielding is $1 - P(Y = 1)$ in Equation (3). For the interaction considered, the outcome will be either vehicle yielding ($Y = 1$) or pedestrian yielding ($Y = 0$). Note that there may be multiple vehicles that do not yield for a particular pedestrian after the first driver, but only the interaction with the first “eligible” driver is considered and it is considered as pedestrian yielding ($Y = 0$). Additional study on the influence of the first driver’s choice to yield or not yield on subsequent drivers may be beneficial but was outside the scope of this study.

$$L(\log(\text{odds})) = \ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n \quad (1)$$

$$P(Y = 1) = \frac{\exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n)}{1 + \exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n)} \quad (2)$$

$$P(Y = 0) = \frac{1}{1 + \exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n)} \quad (3)$$

where β_0 is the constant, β_i is the coefficient of explanatory variable, x_n is the predictor variable, and n is the number of features.

The entire dataset is split into training data and testing data, with 80% of the dataset being used to train the model and 20% of the dataset being reserved as hold-out test data that is used to evaluate the model.

Table 5: Independent Variables for Logistic Regression Models.

FEATURES	ALL VARIABLES	SELECTED VARIABLES	SITE VARIABLES
NUMBER OF PEDESTRIANS	✓		
VEHICLE SPEED	✓	✓	
NUMBER OF BUS STOPS	✓		✓
POSTED SPEED	✓		
CROSSING WIDTH (MAJOR)	✓	✓	
DIST. TO NEAREST PARK	✓	✓	✓
DIST. TO NEAREST SCHOOL	✓	✓	

AADT (MAJOR)	✓		✓
ROAD SURFACE	✓		
PEDESTRIAN TYPE	✓		
OPPOSITE DIRECTION YIELD	✓	✓	
FOLLOWING VEHICLE	✓		
NUMBER OF LANES (MAIN)	✓		✓
WEATHER	✓		
LIGHTING	✓		✓
BIKE LANE	✓		✓
SIGNAGE	✓		✓
MARKINGS	✓		✓
PRESENCE OF SINGLE FAMILY HOUSING	✓		
PRESENCE OF MULTI-FAMILY HOUSING	✓		✓
PRESENCE OF COMMERCIAL	✓		
PRESENCE OF GAS STATION/(CONVENIENT STORE)	✓		
PRESENCE OF RESTAURANTS/BARS	✓	✓	✓
PRESENCE OF PARKING LOTS	✓	✓	✓
PRESENCE OF ON STREET PARKING	✓		
PAWS SCORE	✓		
TREE COVER	✓		✓

5.4.1.1 Model with all variables

A baseline model is constructed using all available data. While some features may prove to not be significant in predicting driver yielding behavior, this model provides a baseline for comparison to other, more sophisticated models.

The first logistic regression model presented in Table 6 is constructed using all the above (significant) variables. The coefficients β_i of the model in Table 6 are estimated using maximum-likelihood estimation, which an iterative search procedure can solve to find values that minimize the error in the predicted probabilities by the model. The coefficients β_i indicate the magnitude of the effects of the corresponding predictor variables on the dependent variable, which is driver yielding. The standard error is an estimate of the standard deviation of the coefficient, and the Z-score describes the deviation from the mean in units of standard deviation. The p-value is a statistical test that shows the probability of the extreme results of the statistical experiment happening. In other words, a low p-value means the predictor is likely to be meaningful to the model.

5.4.1.2 Model with selected variables

Table 6 lists the seven variables that were individually identified as significant and thus included in the logistic regression model of driver yielding. Based on the variable selection criterion introduced above, the vehicle speed, whether a driver in the opposite direction yields, the width of major crossing road, the presence of restaurants/bars/parking lots, the distance to the nearest park, and the distance to the nearest school are all significant and therefore included.

Table 6: Logistic regression model results with selected variables.

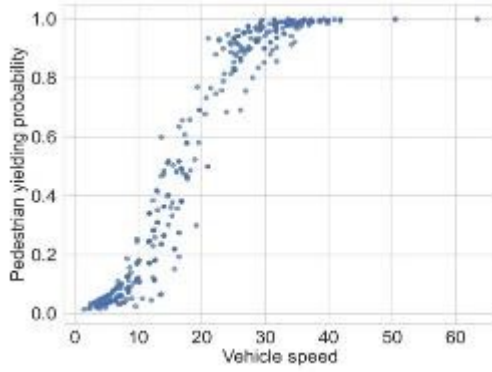
Variables	coefficient	std err	z	P> z 	Effect
Vehicle speed	-0.2354	0.011	-22.357	0.000	negative
Opposite direction yield	-1.1549	0.114	-10.092	0.000	negative
Crossing width (major)	0.1006	0.010	10.154	0.000	positive
Presence of Restaurants/Bars	1.8636	0.382	4.929	0.000	positive
Presence of parking lots	-1.5832	0.253	-6.25	0.000	negative
Dist. to nearest park	-3.53	0.935	-3.774	0.000	negative
Dist. to nearest school	-0.4922	0.143	-3.433	0.001	negative

Based on the result in Table 6, all the selected variables remain significantly important at the 95% confidence level when combined in the model. The odds ratios for all the significant variables are also calculated.

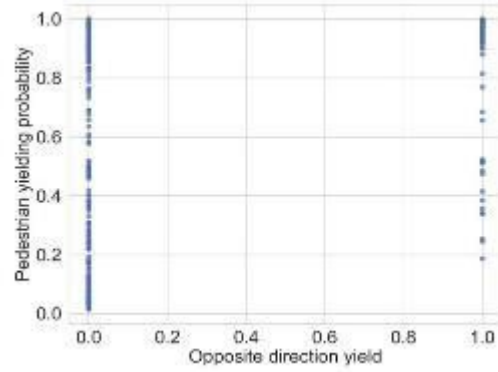
Figure 12 describes how the model predicts the probability of driver yielding with each of the independent variables. Based on the model results, vehicle speed appears to be the most significant variable. The faster the vehicle, the less likely the driver will yield to the pedestrian, as shown in Figure 12a. Vehicle speed has much more variability between interaction events compared with other predictor variables (many of which are binary) in Figure 12.

Note that in Figure 12, some features are distributed along the horizontal (variable) axis while others are concentrated on either extreme. This is a result of different data types, where continuous variables are distributed throughout the variable range, while discrete variables can take only a select few values.

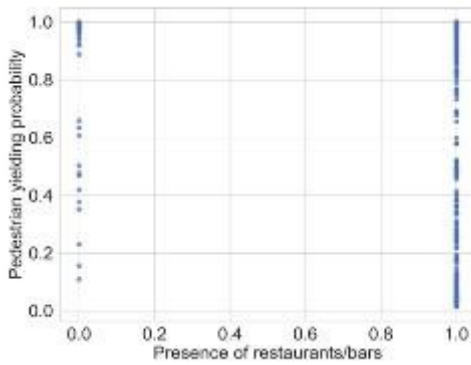
Based on the analysis of statistically significant features, vehicle speed, whether or not a vehicle in the opposing direction yields, the presence of street-adjacent parking lots, an increased distance to the nearest park, and an increased distance to the nearest school all decrease the probability that an observed driver will yield. The crossing width and adjacent land use with bars and restaurants tends to increase the probability that a given observed driver will yield.



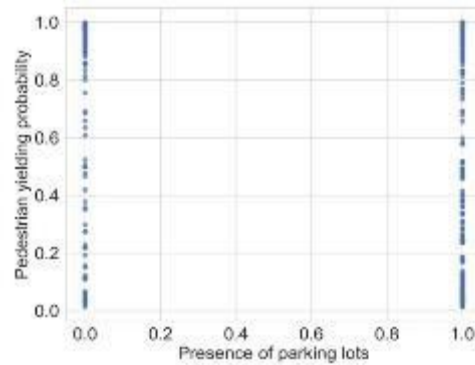
(a) Vehicle speed vs. Driver yielding probability.



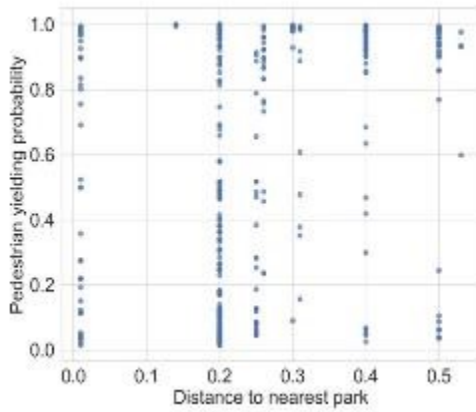
(b) Opposite yielding vs. Driver yielding probability.



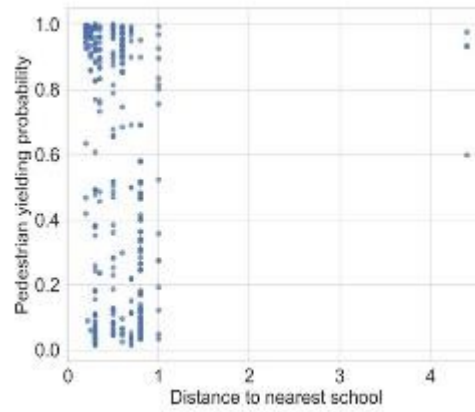
(c) Presence of restaurants/bars vs. Driver



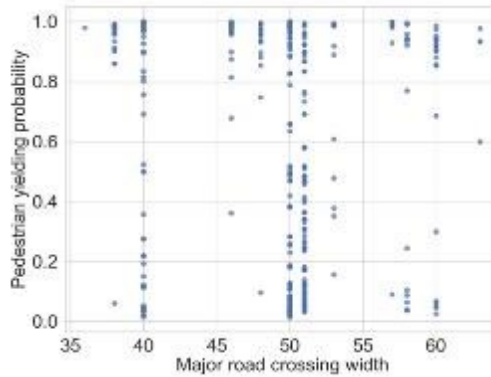
(d) Presence of parking lots. vs. Driver yielding probability.
probability.



(e) Distance to nearest park (mi) vs. Driver yielding



(f) Distance to nearest school vs. Driver yielding probability



(g) Major road crossing width (ft) vs. yielding probability

Figure 12: Predictors vs. yielding probability. In the context of binary variables (shown in ((b), (c), and (d)), 0 means that the feature is not present, while 1 indicates that the feature is present at the site.

5.4.1.3 Model with Site Features

While the analysis in Section 5.4.1.2 provides valuable insight into which features correlate with higher driver yielding rates, the results indicate that many features that cannot be influenced by intersection design are significant. Therefore, to understand what site specific features are relevant for driver yielding, we also constructed a model with selected site features. The model result with only site features (13 variables) are presented in Table 7, and the variables are all statistically significant site variables.

The variables are grouped into two categories based on their effects on driver yielding probability. In this site model, the presence of crosswalk signs, the presence of restaurants/bars, the presence of parking lots, greater distance to a park, more tree coverage, more bus stops, higher AADT, and continental markings have positive influences on driver yielding to pedestrians. In comparison, more lanes on the major road, the presence of bike lanes, the presence of multi-family housing, and the presence of lighting negatively affect driver yielding. The specific impact of each of these features can be computed from the odds ration presented in Table 7.

Table 7: Logistic regression model results with site variables.

Variables	coefficient	std err	z	P> z 	Effect
Number of lanes	-1.8108	0.331	-5.466	0.000	negative
Bike lanes	-2.7107	0.660	-4.109	0.000	negative
Signage	3.6653	0.654	5.604	0.000	positive
Presence of multi-family housing	-4.4568	1.021	-4.367	0.000	
Presence of Restaurants/Bars	3.5323	0.927	3.81	0.000	negative
					positive
Presence of parking lots	2.8063	0.599	4.682	0.000	positive
Dist. to nearest park	5.2798	1.681	3.141	0.002	positive
Tree cover	1.0268	0.332	3.091	0.002	positive
Lighting	-0.9819	0.306	-3.206	0.001	negative
Number of bus stops	0.5076	0.106	4.781	0.000	positive
					positive
Major AADT	0.0001	0.0000289	3.603	0.000	positive
Markings (standard)	-2.5216	0.806	-3.128	0.002	negative
					negative
Markings (unmarked)	-1.6573	0.307	-5.399	0.000	negative

5.4.2 Model Evaluation

The comparison of the three logistic models is presented in Table 5. The results show that the model with selected variables performs best among the three models, and this model uses much fewer variables. Note that all the models have more false positives (Type I Error) than false negatives (Type II

Error) and also have higher recall values than precision. Thus, the model makes more false predictions on driver yielding where in reality the pedestrian yields.

The precision and recall are defined in Equation (4) and Equation (5), respectively. The precision is the number of correctly predicted positive instances divided by the number of total positive predictions the model makes, and the recall is the number of correctly predicted positive instances divided by the total positive instances in the model.

The Receiver Operating Characteristic (ROC) curves are shown in Figure 13. The areas under the ROC curve are 0.88 for the second model, which follows the left-hand border and the top border of the ROC space. This indicates that the model has good performance and balances sensitivity and specificity, which also tell us that the selected variables are essential factors in predicting the outcome of a driver-pedestrian interaction, and provides confidence that the developed models accurately estimate driver yielding.

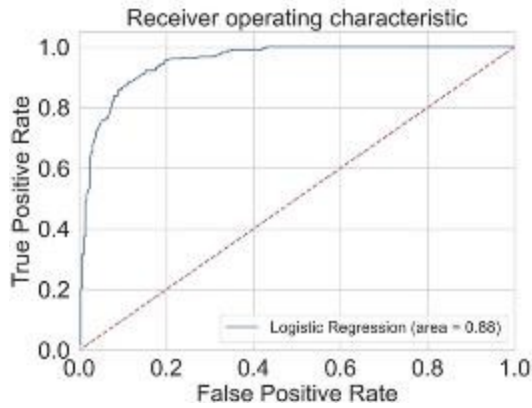
$$\textit{Precision} = \frac{\textit{True positive}}{\textit{True positive} + \textit{False positive}} \quad (4)$$

$$\textit{Recall} = \frac{\textit{True positive}}{\textit{True positive} + \textit{False negative}} \quad (5)$$

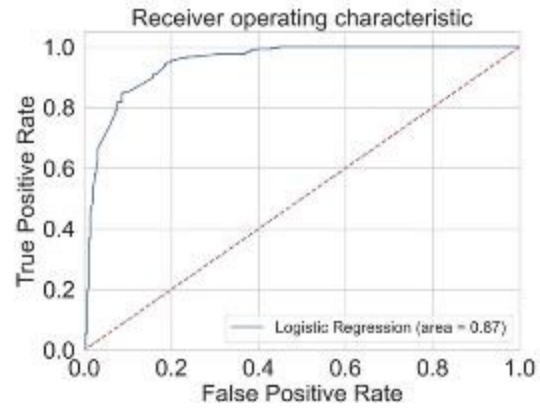
$$\textit{Accuracy} = \frac{\textit{True positive} + \textit{True negative}}{\textit{True positive} + \textit{False positive} + \textit{True negative} + \textit{False negative}} \quad (6)$$

Table 8: Model performance in test data.

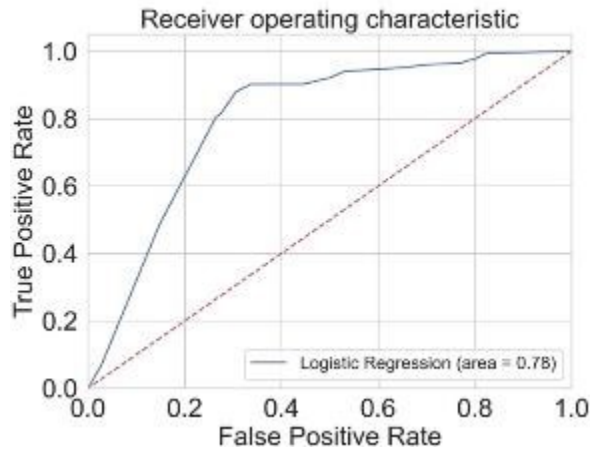
Models	Actual	Predicted	
		Pedestrian yield	Vehicle yield
Model with all features	Pedestrian yield	235	56
	Vehicle yield	19	288
	Accuracy Score	0.87	
	Precision	0.84	
	Recall	0.94	
Model with selected features	Pedestrian yield	237	54
	Vehicle yield	19	288
	Accuracy Score	0.88	
	Precision	0.84	
	Recall	0.94	
Model with site features	Pedestrian yield	193	98
	Vehicle yield	30	277
	Accuracy Score	0.79	
	Precision	0.74	
	Recall	0.90	



(a) Model with all features



(b) Model with selected (significant) features



(c) Model with only site features

Figure 13: ROC curves for different models.

5.4.3 Other models tested

In addition to the logistic regression models presented above, other model forms were also tested. These include k nearest neighbors (kNN), support vector machines (SVM), and neural networks (NN). However, of all four model forms tested, the logistic regression models consistently outperformed the other models in predicting the driver yielding rate.

CHAPTER 6: FINDINGS

6.1 INTERPRETING MODEL RESULTS

In this section, we present how the model results can be used for analysis, and discuss what the data reveal about factors that influence driver yielding.

6.1.1 Interpreting odds ratio

The odds ratio represents the change in probability of an explanatory variable influencing the difference in the response variable. Thus, using the odds ratios presented in Tables 2, 3, and 4, it is possible to compute the sensitivity of the response (yielding) to each variable used in the regression. Specifically, the odds ratio is computed as: $\exp(-\beta_i)$ for each variable i . This tells how much the probability of a driver yielding will change for each unit of change in the predictor variable i .

6.1.2 Model with selected features

Looking at the model with the selected features (both site features and event features), the odds of driver yielding are $\exp(-0.2354) = 0.79$ times lower with each unit (MPH) increase in vehicle speed. Specifically, each unit increase in vehicle speed is associated with a 21% ($1 - 0.79 = 0.21$) reduction in the relative probability of driver yielding, assuming other variables remain fixed (refer to Equation (1)).

Overall, the results show that adjacent land use context is an important factor in yielding. Specifically, the results show that the odds ratio for the presence of restaurants/bars is 6.44 relative to not present, and 0.205 for the presence of parking lot than not presence. Thus, a site that has restaurants/bars has a 544% ($6.44 - 1 = 5.44$) greater relative driver yielding probability than a comparable site that does not have restaurants or bars. This coefficient is large because most of the sites collected in this study have restaurants or bars nearby (see Figures 3) so this variable is skewed. Additionally, the presence of parking lots has a 79.5% ($1 - 0.205 = 0.795$) reduction of relative driver yielding probability compared with the no parking lots group. Longer distances to the nearest park and school negatively affect driver yielding, which has the reductions in the relative probability of 97% and 38% in driver yielding rate with each mile increase.

Specific to the intersection, a wider crossing width² of the major road has a positive influence on driver yielding (10.6% increase in driver yielding probability with each foot of increase in width), which is a relatively small impact, but counters conventional wisdom that wider intersections may have a lower yield rate. However, while found to be statistically significant, this increasing in yielding rate may simply be a statistical anomaly, and a larger sample size may be required in future analysis to conclusively determine the impact of crossing width on yielding behavior. For example, one confounding factor may

²Crossing widths at the 18 sites ranged from 36 ft to 63 ft.

be that, within the sites evaluated, the adjacent land use context at wider roads may be more conducive to observing high rates of driver yielding.

Interestingly, the opposite direction of yielding has an adverse effect on driver yielding (68.5% reduction in relative driver yielding probability compared to if the opposite direction driver does not yield), which is contrary to many studies [9] [4].

However, note that some independent variables have a large standard error and a large p-value. This may be a result of imbalanced data in this study. Specifically, we collected data from 18 locations, but the total number of events varies among the locations, indicating that some intersections may disproportionately influence the results.

The odds ratio for each factor and corresponding units are presented in Table 9 below. Note that since this is the best-fit model, it is the one used to compute the odds ratios. For variables with true/false units, this indicates that the variable represents whether or not that infrastructure feature is present. For example, for Presence of restaurants/bars, the variable indicates whether or not there are restaurants or bars present.

Table 9: Odds ratio for different factors indicating change in driver yielding rate for each unit of change in factor.

Variables	Percent change (%)	Units
Vehicle speed	-21%	MPH
Opposite direction yield	-68%	True/false
Crossing width (major)	11%	Feet
Presence of Restaurants/Bars	545%	True/false
Presence of parking lots	-79%	True/false
Dist. to nearest park	-97%	Miles
Dist. to nearest school	-39%	Miles

6.1.3 Model with site features

A model with only site features was created based on the assertion that designers and engineers may have control over some of the site features (e.g., the presence of signs) but generally do not have control over event features (e.g., pedestrian group size). Since the model with site features only has a lower predictive power than the model with the selected features above, computing the odds ratio in this case may lead to an inaccurate assessment of impact of each variable. Instead, it's important to note which features positively influence yielding, and which negatively influence yielding as seen in Table 10.

Generally, factors that decrease the chance of yielding most significantly are the number of lanes, the presence of bike lanes, and the presence of multi-family housing nearby while the factors that have the most significant positive impact on driver yielding are the presence of signs for the crosswalk, the presence of trees, and the proximity to the nearest park. These results are generally consistent with the best-fit model with both site and event features (the model with the selected features).

6.2 DISCUSSION OF FACTORS THAT IMPACT YIELDING

Based on the model with site and event features, each site feature is listed in Table 10 with the corresponding odds ratio. The odds ratio measures how significantly this factor impacts yielding. I.e., for each 1 unit of change in the quantity, the odds ratio tells the percent change in the yield rate. Note that for binary features such as the presence of signs, the odds ratio tells how much the presence of signs at the crosswalk will increase driver yielding compared to the baseline of no signs for the crosswalk. The identified factors in Table 10 are ranked in order of how much they increase the driver yield response. Thus, factors near that top provide the most significant increase in yielding rates, while those at the bottom are found to correlate with the most significant decrease in driver yielding. The complete list of factors considered is included in the Task 8 deliverable.

Table 10: Logistic regression model results with site variables.

Variables	Coefficient	Odds ratio (%)	Std Err	z	P> z	Effect
Presence of signs	1.2018	232	0.208	5.789	0.000	Positive
Presence of Restaurants/Bars	1.131	210	0.348	3.25	0.001	Positive
Presence of parking lots	0.4173	52	0.155	2.69	0.007	Positive
Number of bus stops	0.1877	21	0.036	5.197	0.000	Positive
Number of lanes	-0.4256	-35	0.078	-5.435	0.000	Negative
Bike lanes	-0.911	-60	0.287	-3.169	0.002	Negative
Presence of multi-family housing	-1.0559	-65	0.368	-2.873	0.004	Negative
No markings	-1.377	-75	0.163	-8.428	0.000	Negative

6.2.1 Significance of signs on driver yielding

Signs indicating the presence of the crosswalk are found to be the most significant site feature factor in determining driver yielding. Based on the model, drivers are about twice as likely to yield at intersections with signs compared to similar intersections that do not have signs at the crossing. This is clearly seen in Figure 14 where the average driver yield rate for intersections both with and without signs at the crossing are compared. The average yield rate for drivers at intersections without a sign is 32.78% while the average driver yield rate at intersections with a sign marking the crossing is 66.51%. Of the site features, this is found to be the biggest indicator of whether or not a driver is likely to yield at an intersection.

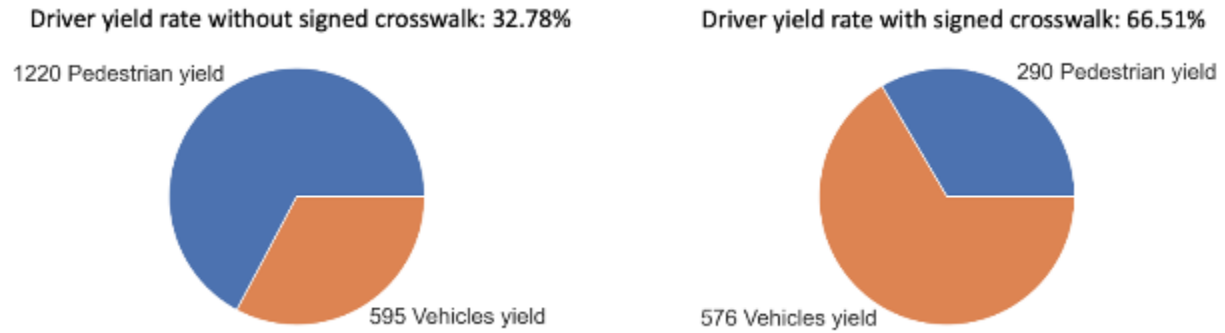


Figure 14: Impact of a sign marking the presence of a pedestrian crossing on the driver yield rate.

6.2.2 Significance of adjacent land use on driver yielding

The adjacent land use is found to be a significant indicators of driver yielding. Of those evaluated, the two most significant indicators are the presence of restaurants and bars (14 of the 18 sites) and the presence of parking lots (10 of the 18 sites) near the pedestrian crossing. For the purpose of this analysis, a one block radius (i.e., to the nearest intersection in each cardinal direction) was used as the proximity for the presence of adjacent land use factors. It is worth noting that such factors may be significant further from the intersection in question along the approach path. However, these were not considered in this study.

Both the presence of bars or restaurants near the intersection as well as the presence of parking lots were found to increase driver yielding rates. Specifically, areas with bars or restaurants saw a 210% higher yielding rate than comparable intersections without bars or restaurants while intersections with adjacent parking saw a 52% higher yielding rate over comparable intersections without parking in a close proximity.

Another land use factor found to influence the yielding rate is the presence of multi-family housing in close proximity (within a one block radius) of the intersection. These crossings are found to generally have lower driver yielding rates (65% reduction) than intersections without multi-family housing buildings. This is likely a result of the built environment around multi-family housing buildings.

These results don't imply that either bars/restaurants cause a higher yielding rate, or that multi-family housing causes a lower yielding rate. However, their significance indicates that the built environment may play some role in driver yielding rates. These can thus be thought of as proxy measures for a more difficult to define concept of an urban built environment that leads to high yielding rates. Importantly, the presence or absence of any of these features alone does not necessarily change the yielding rate. Instead, these are simplistic factors that can be used to approximate the adjacent land use and how the driver may perceive the intersection with respect to their yielding decision.

Adjacent land uses that did not have a statistically significant impact on driver yielding (positive or negative) included the presence of single-family housing, commercial buildings, or gas station/convenience store.

6.2.3 Significance of adjacent bus stops on driver yielding

Similar to the adjacent land use context features discussed in Section 2.2, the presence of bus stops (11 of the 18 sites) within a 1 block radius of the crossing lead to higher yielding rates when compared to a similar intersection without bus stops. For each additional bus stop within one block of the crossing, the driver yielding rate was found to increase by 21%.

6.2.4 Significance of the number of lanes on driver yielding

The number of lanes is found to be a significant factor in determining driver yielding with roads with more lanes leading to lower yielding rates. For each additional lane that an intersection has, yielding is expected to decrease by roughly 35% compared to a similar intersection with fewer lanes. This impact is clearly seen in Figure 2 where the average yield rate across all intersections where the major road has two lanes is 53.07%. This yield rate decreases to 13.12% for larger roads with 4 lanes.

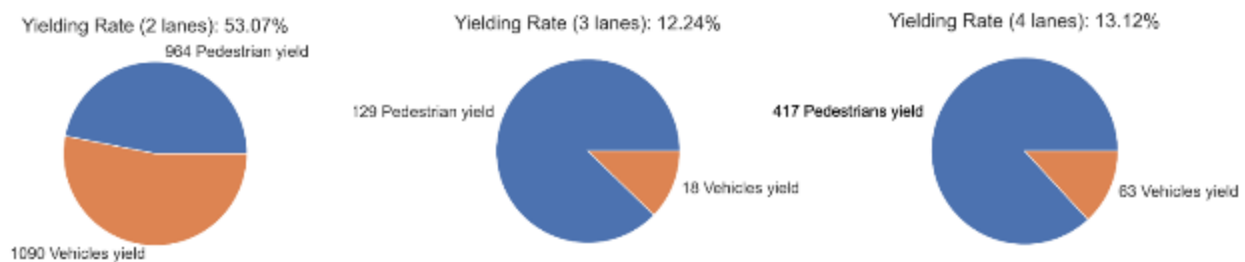


Figure 15: Impact of number of lanes on driver yield rate.

6.2.5 Significance of bike lanes on driver yielding

Roads with bike lanes are found to generally have lower yielding rates than roads without bike lanes. Specifically, roads with bike lanes are found to have a 60% lower driver yielding rate compared to similar roads without bike lanes. This may result from roads with bike lanes generally being larger, higher traffic volume roads where undesirable conditions exist for cyclists without the presence of bike lanes. Therefore, this finding may reflect a broader finding about the land use context that an intersection is in, and does not indicate that bike lanes will specifically reduce driver yielding rates.

6.2.6 Significance of markings on driver yielding

Intersections without crosswalk markings (12 of the 18 sites) are found to have significantly lower driver yielding rates than intersections that are marked (6 of the 18 sites) with either continental (horizontal bars) or standard (parallel line) markings. When comparing intersections without markings, drivers are 75% less likely to yield than at comparable intersections with markings.

As seen in Figure 16 below, the driver yield rate is lowest (10.81%) for unmarked intersections, while the driver yield rate is highest for crossings with standard markings, where a total of 69.27% of drivers yield to pedestrians. When comparing standard and continental markings, standard markings result in the highest driver yield rate, though other factors may be confounding this finding. For example, in the specific intersections identified, there may be other factors leading to particularly high yield rates at crossings with standard markings.

While Figure 16 below is not able to normalize for other influencing factors, these factors are normalized for in the analysis presented in Table 10, which finds that unmarked crossings have the most significant adverse impact on driver yielding rates.



Figure 16: Comparison of driver yield rates for different types of crossing markings.

6.2.7 Other significant nonsite-specific factors in driver yielding

All the significant features identified in Table 2 and discussed above are site-specific features meaning that they relate to the conditions at a site, and do not vary between different driver-pedestrian interactions at the site. In addition to these features, some interaction-specific features are also identified as being associated with higher driver yield rates. While these features may be outside the scope of what transportation agencies can influence, knowledge of these features and their impact on driver yielding may be informative in guiding design decisions. Therefore, a brief discussion of these features and their impact on yielding is presented below.

Party size (i.e., number of pedestrians trying to cross the road at the same time) is found to significantly impact driver yielding, with larger parties resulting in a higher driver yield rate, as shown in Figure 4 below. For example, when a single pedestrian is trying to cross the street, the average yield rate across all intersections is 37.39% while the driver yield rate for parties of 5 pedestrians is 70%.

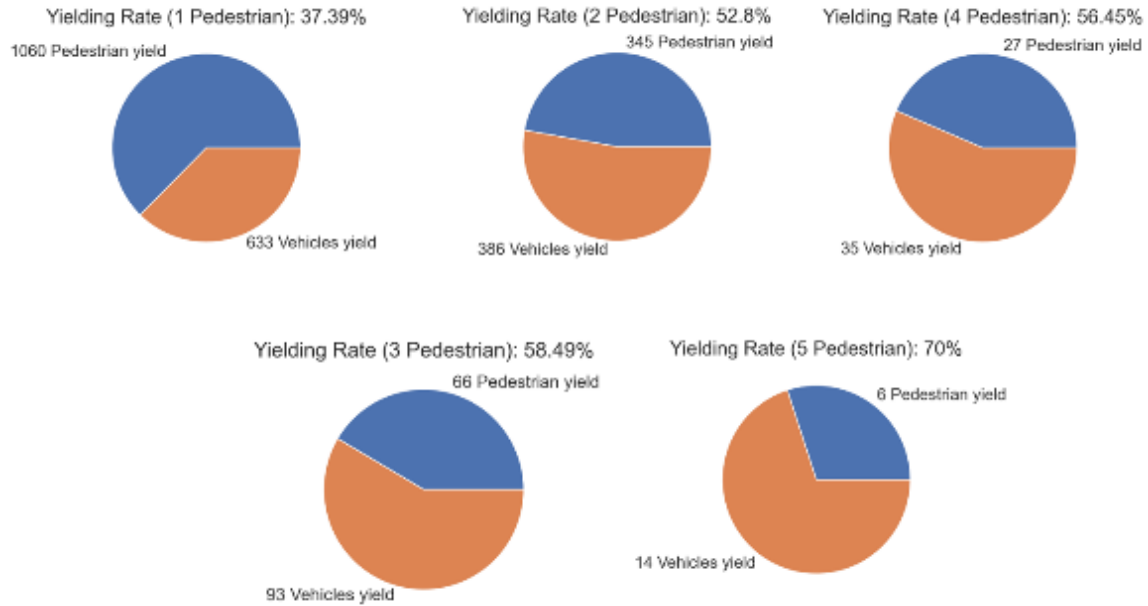


Figure 17: Impact of party size on driver yielding rate. Larger party sizes see a significantly higher driver yield rate.

Another factor found to be significant on driver yielding is the vehicle type, since not all drivers are found to be equally likely to yield, and the type of vehicle they drive impacts the average driver yield rate. As seen in Figure 5, passenger vehicles generally have the lowest driver yield rate with sedans yielding 40.98% of the time and SUVs yielding 43.81% of the time. Vehicles often driven by professional drivers generally exhibit higher driver yield rates, with busses yielding 65.08% of the time.

While these factors are beyond the control of local road management agencies, knowing these interaction-specific trends can help guide design decisions, especially if certain trends are observed locally.

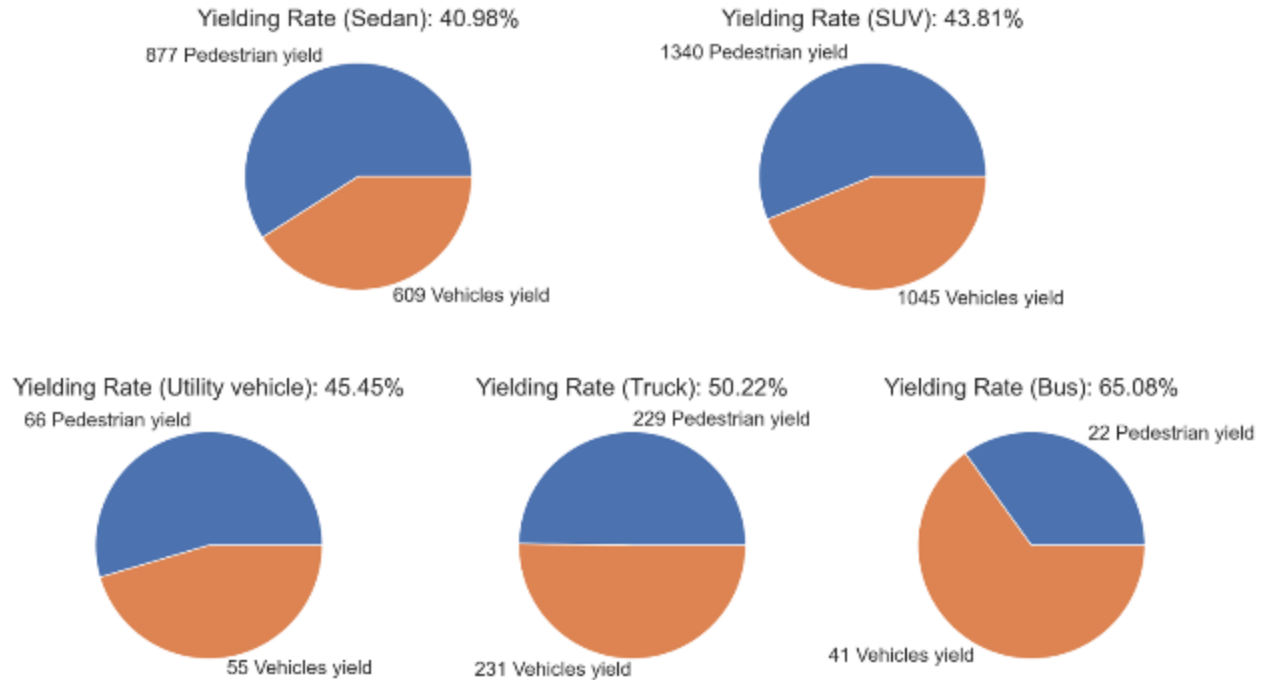


Figure 18: Impact of vehicle type on driver yield rate.

6.2.8 Relationship between crash history and driver yielding rates

An brief analysis was also conducted to understand how well the crash history at an intersection predicts the driver yielding rate. The number of crashes within a 300 foot radius of each crossing were counted at each intersection within the past five years. The analysis to correlate driver yielding rate at each intersection with the number of crashes indicates that there was no correlation between this crash metric and driver yielding (Figure 6). Note that this only considers the total number of crashes, and not the crash rate or the severity of the crash.

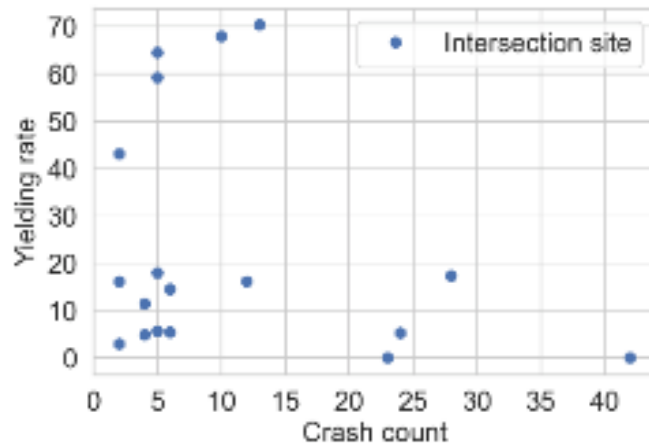


Figure 19: Comparison of crashes and driver yield rate showing no significant correlation between the two.

This suggests that, while crash data is important from a safety perspective and is a useful tool for assessing safety-targeted interventions, a retrospective crash analysis may not provide sufficient insight for understanding the level of service pedestrians experience with respect to driver yielding rates.

CHAPTER 7: RECOMMENDATIONS AND CONCLUSIONS

7.1 DESIGN RECOMMENDATIONS FOR INCREASING DRIVER YIELDING

Based on the odds ratios identified by the model and the data presented above, the following recommendations are made to encourage driver yielding at unsignalized intersections where driver yielding is desired.

Crash Data: While historical crash data has been found to be a good predictor of future crashes and is an important tool for improving roadway safety, high-level crash data was not found to be sufficiently informative when considering driver yielding. Therefore, it was recommended that additional data such as geometric features of the road be considered to increase driver yielding rates, but further analysis may be needed to consider specific crash types such as rear end collisions.

Number of lanes: Generally, road width should be as small as possible to safely accommodate traffic volume. Each additional lane is found to decrease driver yielding by roughly 35%. This is likely because roads with more lanes generally are wider. Reducing roadway widths will likely increase driver yielding and should be done in balance with maintaining sufficient roadway capacity.

Crossing Markings: Pedestrian crossing markings should be included in crossing designs at sites where high driver yield rates are desired. Both continental and standard markings are correlated with higher yielding rates when compared to unmarked intersections.

Signs for crossing: Signs indicating the presence and location of a crossing should be included in the design at sites where high driver yielding rates are desired. The presence of a sign was found to more than double the probability that a driver will yield to a pedestrian.

Presence of bike lanes: Intersections with bike lanes are generally found to have lower yielding rates than intersections without bike lanes. However, it is important to note that bike lanes inherently increase the intersection width, which may be the leading factor for this finding. Therefore, additional data is needed to conclusively determine the impact of bike lanes specifically on driver yielding.

Adjacent land-use context: While the adjacent land-use context may not be a feature that can be altered by a transportation agency, the impact the land-use context has on driver yielding behavior is substantial, and the adjacent land use should be considered when designing intersections. Based on the findings, intersections in neighborhoods with a high restaurant density have higher yielding rates, while the presence of multi-family housing tends to be correlated with lower driver yielding rates. While these findings do not imply causation, they are worth considering when designing intersections. The reader is referred to the brief discussion on the difference between correlation and causation in the introduction.

Operating speed: Operating speeds above 25 mph were shown to strongly correlate with low driver yielding rates. Roadway features that reduce driver operating speed should be considered at sites where high driver yielding rates are desired.

REFERENCES

- [1] National Highway Transportation Safety Administration (NHTSA). (2020) .
- [2] B. F. Herms. (1972). Pedestrian crosswalk study: Accidents in painted and unpainted crosswalks. *Highway Research Record*, 406, 1–13.
- [3] K. Ralph, E. Iacobucci, C. G. Thigpen, & T. Goddard. (2019). Editorial patterns in bicyclist and pedestrian crash reporting. *Transportation Research Record*, 2673, 663–671.
- [4] D. Akin & B. Akba. (2010). A neural network (NN) model to predict intersection crashes based upon driver, vehicle and roadway surface characteristics. *Scientific Research and Essays*, 5, 2837–2847.
- [5] I. Banerjee & D. R. Ragland. (2007). Evaluation of countermeasures: A study on the effect of impactable yield signs installed at intersections in San Francisco. Transportation Research Board.
- [6] Y. Zheng, T. Chase, L. Elefteriadou, B. Schroeder, & V. P. Sisiopiku. (2015). Modeling vehicle–pedestrian interactions outside of crosswalks. *Simulation Modelling Practice and Theory*, 59, 89–101.
- [7] J. Zhao, J. O. Malenje, J. Wu, & R. Ma. (2020). Modeling the interaction between vehicle yielding and pedestrian crossing behavior at unsignalized midblock crosswalks. *Transportation Research Part F: Traffic Psychology and Behavior*, 73, 222–235.
- [8] D. Sun, S. V. S. K. Ukkusuri, R. F. Benekohal, & S. T. Waller. (2003). Modeling of motorist-pedestrian interaction at uncontrolled mid-block crosswalks. *Transportation Research Record, TRB Annual Meeting CD-ROM*.
- [9] B. J. Schroeder & N. M. Roupail. (2011). Event-based modeling of driver yielding behavior at unsignalized crosswalks. *Journal of Transportation Engineering*, 137, 455–465.
- [10] K. Salamati, B. J. Schroeder, D. R. Geruschat, & N. M. Roupail. (2013). Event-based modeling of driver yielding behavior to pedestrians at two-lane roundabout approaches. *Transportation Research Record*, 2389, 1–11.
- [11] A. P. Silvano, H. N. Koutsopoulos & X. Ma. (2016). Analysis of vehicle-bicycle interactions at unsignalized crossings: A probabilistic approach and application. *Accident Analysis & Prevention*, 97, 38–48.

- [12] S. Turner, K. Fitzpatrick, M. Brewer & E. S. Park. (2006). Motorist yielding to pedestrians at unsignalized intersections: Findings from a national study on improving pedestrian safety. *Transportation Research Record*, 1982, 1–12.
- [13] D. Muley, M. Kharbeche, W. Alhajyaseen & M. Al-Salem. (2017). Pedestrians' crossing behavior at marked crosswalks on channelized right-turn lanes at intersections. *Procedia Computer Science*, 109, 233–240.
- [14] L. Bai, P. Liu, Y. Chen, X. Zhang & W. Wang (2013). Comparative analysis of the safety effects of electric bikes at signalized intersections. *Transportation Research Part D: Transport and Environment*, 20, 48–54.
- [15] R. J. Schneider, M. C. Diogenes, L. S. Arnold, V. Attaset, J. Griswold & D. R. Ragland. (2010). Association between roadway intersection characteristics and pedestrian crash risk in Alameda County, California. *Transportation Research Record*, 2198, 41–51.
- [16] R. Van Houten & L. Malenfant. (1992). The influence of signs prompting motorists to yield before marked crosswalks on motor vehicle-pedestrian conflicts at crosswalks with flashing amber. *Accident Analysis & Prevention*, 24, 217–225.
- [17] R. O. Phillips, T. Bjørnskau, R. Hagman, & F. Sagberg, (2011). Reduction in car–bicycle conflict at a road–cycle path intersection: Evidence of road user adaptation? *Transportation Research Part F: Traffic Psychology and Behavior*, 14, 87–95.
- [18] K. J. Krizec. (2007). Estimating the economic benefits of bicycling and bicycle facilities: An interpretive review and proposed methods. In *Essays on transport economics* (pp. 219–248.) New York: Springer.
- [19] P. L. Jacobsen. (2015). Safety in numbers: More walkers and bicyclists, safer walking and bicycling. *Injury Prevention*, 1, 271–275.
- [20] W. K. M. Alhajyaseen, M. Asano, & H. Nakamura. (2012). Estimation of left-turning vehicle maneuvers for the assessment of pedestrian safety at intersections. *IATSS Research*, 36, 66–74.
- [21] H. Iasmin, A. Kojima & H. Kubota. (2015). Yielding behavior of left turning driver towards pedestrian/cyclist: Impact of intersection angle. *Journal of the Eastern Asia Society for Transportation Studies*, 11, 2146–2158.
- [22] J. Stipancic, S. Zangenehpour, L. Miranda-Moreno, N. Saunier, & M.-A. Granié. (2016). Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. *Accident Analysis & Prevention*, 97, 19–27.

- [23] X. Zhang, S. T. Waller & P. Jiang. (2020). An ensemble machine learning-based modeling framework for analysis of traffic crash frequency. *Computer-Aided Civil and Infrastructure Engineering*, 35, 258–276..
- [24] M. K. D. G. Kleinbaum. (2002). *Logistic regression: A self-learning text*. New York: Springer.
- [25] L. Rodegerdts, M. Blogg, E. Wemple, E. Myers, M. Kyte, M. Dixon, G. ... & D. Carter. (2007). Appendixes to NCHRP Report 572: Roundabouts in the United States. *NCHRP Web-Only Document*, 94, 1-115.
- [26] K. Fitzpatrick, S. M. Turner, M. Brewer, P. J. Carlson, B. Ullman, N. D. Trout, E. S. ... & D. Lord. (2006). Improving pedestrian safety at unsignalized crossings: Appendices B to O. Transportation Research Board.
- [27] P. Morency, L. Gauvin, C. Plante, M. Fournier, & C. Morency. (2012). Neighborhood social inequalities in road traffic injuries: The influence of traffic volume and road design. *American Journal of Public Health*, 102, 1112–1119.
- [28] 2. Highway Capacity Manual, 2010.
- [29] R. Elvik. (2009). The non-linearity of risk and the promotion of environmentally sustainable transport. *Accident Analysis & Prevention*, 41, 849–855.
- [30] D. S. Pawar & G. R. Patil. (2016). Critical gap estimation for pedestrians at uncontrolled mid-block crossings on high-speed arterials. *Safety Science*, 86, 295–303.
- [31] L. D. Frank & P. O. Engelke. (2001). The built environment and human activity patterns: Exploring the impacts of urban form on public health. *Journal of Planning Literature*, 16, 202–218.
- [32] J. Pucher & R. Buehler. (2017), *Cycling towards a more sustainable transport future*, Milton Park, UK: Taylor & Francis.
- [33] J. Hourdos, P. Dirks, D. Lehrke, G. Parikh, G. Davism & C. Cheong. (2020). Assessing the impact of pedestrian-activated crossing systems. Minnesota Department of Transportation.
- [34] R. Van Houten & C. Seiderman. (2005). How pavement markings influence bicycle and motor vehicle positioning: Case study in Cambridge, Massachusetts. *Transportation Research Record*, 1939, 2–14.

- [35] R. L. Sanders. (2016). We can all get along: The alignment of driver and bicyclist roadway design preferences in the San Francisco Bay Area. *Transportation Research Part A: Policy and Practice*, 91, 20–133.
- [36] I. of Transportation Engineers, 1982.
- [37] B. C. Langford, J. Chen, & C. R. Cherry. (2015). Risky riding: Naturalistic methods comparing safety behavior from conventional bicycle riders and electric bike riders. *Accident Analysis & Prevention*, 82, 220–226.
- [38] R. Gang & Z. Zhuping. (2011). Traffic safety forecasting method by particle swarm optimization and support vector machine. *Expert Systems with Applications*, 38, 10420–10424.
- [39] N. Chintalacheruvu, V. Muthukumar and others. (2012). Video based vehicle detection and its application in intelligent transportation systems. *Journal of Transportation Technologies*, 2, 305.
- [40] R. Aldred. (2016). Cycling near misses: Their frequency, impact, and prevention. *Transportation Research Part A: Policy and Practice*, 90, 69–83.

APPENDIX A: LITERATURE REVIEW

DRIVER YIELDING BEHAVIOR AT INTERSECTIONS

Understanding the factors that lead to driver yielding is crucial for traffic safety analysis and pedestrian safety. A pedestrian-driver interaction is the outcome of a pedestrian arriving at an intersection while a driver is approaching the same intersection. Thus, the approaching driver can yield to the pedestrian (driver yielding) by either stopping or slowing to accommodate the pedestrian, or the driver can also decide to not yield to the pedestrian, resulting in a non-yielding event.

Understanding pedestrian and driver behavior and their interactions are important to design safer streets and reduce pedestrian fatalities. The low compliance rate of drivers yielding to pedestrians at unsignalized intersections has been documented through many studies. A summary of the findings is presented by [25], [26]. Generally, these studies find that in the United States driver yielding rates are quite low, even in marked and signed crosswalks. As a result, pedestrians are often forced to wait until a suitable gap in traffic opens before crossing the street. A theme that is consistent across many different studies is the focus on mid-block crossings (often colloquially referred to as jay-walking). While far more ambiguous from a right-of-way perspective, studies of mid-block crossings are disproportionately represented in the literature, with few studies identified that do not consider mid-block crossings.

Summarizing the prevailing literature, studying driver-pedestrian interactions has been an area of intense interest over the last half century. While some studies look at interactions at intersections, many studies consider mid-block crossings where right-of-way is less well defined. Generally, studies find higher yielding rates at marked crosswalks (and even at unmarked crosswalks at intersections) than for mid-block crossings.

CRASHES AT INTERSECTION

Another source of information related to pedestrian safety and driver yielding is pedestrian crashes. While these don't directly address the factors that lead to driver yielding, a pedestrian crash represents a case where a driver (or pedestrian) failed to yield. Thus, understanding factors that lead to higher crash rates can help reduce risk for drivers. Importantly, while, as shown in Section 2, data is scarce and difficult to collect on driver-pedestrian interactions, ample data has been collected on pedestrian-involved crashes. Thus, with the richness of data, pedestrian crashes have been studied more frequently than driver-pedestrian interactions. This section provides an overview of relevant literature on pedestrian crashes.

Understanding how to model crash frequency has been a topic of significant interest. This is true both for crashes with pedestrians involved as well as for vehicle-only crashes. Important here is the methodology used when constructing the crash model. For example, Akin et al. [4] developed a neural network (NN) model to predict intersection crashes in Macomb County, Michigan. The research team tested 48 independent variables in the NN model to model crash behavior. Data from a total of 16,384 crashes were used to build the model, and the resulting had an accuracy of 90.9% in predicting whether or not a crash would occur given a particular scenario. While this study is not specific to crashes with

pedestrian involvement, the results show that using the correct statistical analysis can result in highly accurate crash modeling.

Understanding how societal factors such as social inequality influences crash risk has also been considered. For example, Morency, et al. [27] investigated the relationship between traffic crash prevalence and social inequalities. They studied how inequalities in wealth of the surrounding area influenced crash risk, and model the crash prevalence via statistical analysis. They found there are significantly more injured pedestrians, bikers, and vehicle occupants at intersections near poorer areas as opposed to areas with higher wealth. While this may be since wealthier areas have higher rates of driving, the authors do not speculate on the cause for this finding.

Another prevalent avenue to approach the topic of crash modeling is to consider what intersection characteristics influence crash risk. For example, Scheirder et al. [15] studied the intersection characteristics associated with pedestrian crash risk and used crash data and pedestrian volumes at 81 intersections in Alameda County, California to conduct this analysis. The pedestrian crash data were taken from the California Highway Patrol Statewide Integrated Traffic Records System database for the 10-year period between 1998 and 2007. The research team built a negative binomial regression model and tested over 30 variables and with a natural logarithm form of some of the exposure variables. The final model had eight explanatory factors that were found to be statistically significant at the 0.05 level. Specifically, these features related to intersection features, vehicle movement, and land use properties. The results showed crash rates unsurprisingly depend on both pedestrian and vehicle volumes. Additionally, the number of right-turn-only lanes at intersections and the number of commercial retail properties also were found to increase the pedestrian crash rate. Finally, pedestrian age was found to be significant, with younger pedestrians more likely to be involved in a crash. These findings show that intersection characteristics are an important factor in crash rates for pedestrians.

While studying and modeling crash likelihood is not directly related to the question of drivers yielding to pedestrians, the literature provides insight into what factors are found to be significant in understanding crashes, which are clearly related to yielding. Thus, as stated by Akin's, et al. [4], while studying crashes is clearly distinct from studying yielding, and there are more factors to consider in crash studies, such as impact location, the overlap in factors found to be significant driver yielding studies is substantial. Thus, this may be fruitful when trying to identify what factors are significant for driver yielding, since more research has been conducted on modeling crashes.

PEDESTRIAN BEHAVIOR AT INTERSECTIONS

In addition to driver behavior, pedestrian behavior at intersections is another commonly studied topic that may be relevant to the research at hand. For example, gap acceptance theory is widely used to evaluate the safe operation of a road at uncontrolled intersections (HCM 2010 [28]). The underlying idea is that pedestrians have to wait for safe gaps to cross at intersections. While waiting to cross a roadway, a pedestrian will either accept or reject each gap that arises. The longer a pedestrian waits, the shorter the gap they're willing to accept. A driver may be forced to react differently if the pedestrian accepts a shorter gap than the driver is expecting. Raff, et al. [17], studied the temporal and spatial gap

acceptance at uncontrolled mid-block street crossings for pedestrians and found that both temporal and spatial gaps follow a lognormal distribution, suggesting that there is structure to pedestrian behavior across the population [28].

From the urban system level, Elvik [29] explored the non-linearity of injury risks for pedestrians and cyclists and found that as more people transfer from motor-vehicle trips to walking or biking, the total number of crashes will be reduced dramatically. This shows that encouraging walking or biking will not necessarily lead to more crashes. Instead, this promotes the idea of “strength in numbers” suggesting that higher pedestrian volumes could actually reduce both the prevalence and the total number of pedestrian crashes.

In summary, pedestrian behavior characteristics play a substantial role in determining the outcome of driver-pedestrian interactions. Importantly, while there are some features that lead to higher yielding rates, it’s important to keep in mind that pedestrians generally have the right of way at crosswalks, and therefore it is important to avoid blaming the pedestrian for engaging in behavior that leads to lower yielding rates. Some interesting findings that summarize the significance of pedestrian behavior include that an increase in walking or biking will not necessarily lead to more crashes [29]. In fact, it may actually lead to a reduction in pedestrian-involved crashes. Other factors such as gender, party size, pedestrian type, and waiting position will also influence the outcome of the driver-pedestrian interaction [30], [13], [22]. These characteristics along with gap acceptance can be modeled using statistical methods [30].

DRIVER BICYCLE INTERACTIONS

While there is only limited research on drivers and pedestrians, an area with many parallels is that of driver-bicycle interactions. This has been extensively studied to understand how to make roadways safer for bicyclists. Therefore, we provide a summary of bike-related literature that we feel is relevant to understanding drivers yielding to pedestrians.

The idea that the urban built environment is influential in cyclists’ decisions has been well established. For example, Frank and Engelke [31] studied how walking and biking are important for public health and how urban factors can influence people’s interest or willingness to participate in these activities. They also suggest rethinking public policy for transportation investment and land development. Pucher and Buehler [32] argued how cycling is important and can lead to a more sustainable future. Figure 7 below shows the increasing bike mode in large cities. Jacobsen [19] studied the relationship between the number of people walking or cycling and the frequency of collisions, and use population level and time-series data to validate the results. The results showed that vehicle drivers adjust their behavior when people are walking and biking, and drivers are less likely to collide when more people walk or bike together. Krizec [18] provides a review of existing literature of estimating the economic aspects and benefits of cycling and bicycle facilities. Krizec [18] also suggests comprehensive methodologies for future work.

In conclusion, while not directly related to pedestrian safety, the area of bicyclist safety is one with a rich literature that has many parallels to that of pedestrian safety. The key findings identified in this

literature review indicate that high-quality infrastructure along with educational interventions will increase the compliance rate and reduce crashes, both with respect to drivers and bicyclists. The literature also identifies factors such as gender, speed, and type of bicycle which influence overall safety. These are factors that may well prove valuable in trying to improve the walking experience for pedestrians as well.

INTERSECTION CHARACTERISTICS

A factor that goes beyond individual driver or pedestrian behavior is the layout and design of an intersection, which has been the subject of several studies. The idea being that certain design features may make transportation infrastructure safer (or less safe) for pedestrians. Understanding which features correlate positively with a safe pedestrian experience will help provide guidelines or modifications to current and future intersections to increase safety and walkability.

One subject that was found repeatedly in the literature is the angle of the two roads at an intersection, and the influence it has on driver yielding. For example in [21] the authors found that the crash data shows that left-turning vehicles did not always give priority to the pedestrians and that in particular, drivers at obtuse-angled intersections were less likely to yield to pedestrians and tended to accept smaller gaps and turn quickly than at right-angled intersections. Drivers at acute-angled intersections were more likely to yield since they were very careful about turning and slowed more substantially. This can be thought of as a proxy for the drivers' confidence in making the turn, with lower driver confidence leading to higher yielding rates to pedestrians.

Similarly, Alhajyaseen, et al. [20] proposed four empirical models to represent the variations in the maneuvers of left-turners at signalized intersections, which considers the stochastic characteristics of driver behavior and the effects of intersection layout. Validated results from a Monte Carlo simulation showed that the proposed model can represent the variations statistically. The sensitivity analysis showed that vehicle speed is related to the intersection geometric characteristics such as corner radius, intersection angle, and crosswalk position. The results also showed that conflicts are less severe at compact intersections while having a higher possibility of collisions.

Additionally, some studies investigate different engineering treatment such as signals and marked crosswalks on the yielding behavior. For example, Turner, et al. [12] collected data from 42 study sites in different regions across the country to evaluate the effectiveness of different engineering treatments to improve the safety of pedestrians crossing in marked crosswalks. The research team categorized the crossing treatments into three basic categories: 1) Red signal or beacon, devices that display a circular red indication to drivers; 2) Active when present, devices that are designed to display a warning only when pedestrians are present or crossing the street; 3) Enhanced or high-visibility, devices and design treatments that enhance the ability of pedestrian to cross the street or the visibility of the crossing locations. They found that the red signal or beacon treatments perform best with compliance rates greater than 94%, while the last two types of treatment have a wide range of compliance rates. Additionally, street characteristic variables such as the number of lanes and speed limit are all significant variables at 0.05 level and 0.1 level from the data analysis. In contrast, Hourdos, et al [33] found that

pedestrian-activated signals were an effective way to ensure pedestrians are seen by drivers. Differing results in different studies and suggesting that other factors may also be involved.

Van Houten and Malenfant [34] studied the influence of signs prompting drivers to yield as well as vehicle-pedestrian conflicts at crosswalks. They chose two sites located in Dartmouth and Nova Scotia, Canada, where a trained observer collected data and scored the behavior each weekday. There are four types of experiment conditions, (i) Baseline with the absence of the "STOP HERE FOR PEDESTRIANS" sign; (ii) Sign alone, the "STOP HERE FOR PEDESTRIANS" sign was erected about 50 feet before the crosswalk; (iii) Sign plus stop line, the sign remained in place and the advance lines were placed; and (iv) Follow-up data were collected one month and one year after the termination of the experiment.

Van Houten and Malenfant [34] found that the presence of a sign will increase the distance that drivers stop before crosswalks when yielding to pedestrians, and the addition of stop lines increases the distance and further reduces the percentage of motor vehicle-pedestrian conflicts as well. However, they also found that only 59% of the pedestrians activated the light on Portland Street, and only 28% activated the light on Prince Albert Road where the traffic was relatively less busy. It should be noted drivers were only likely to yield when the crosswalk light was activated. These results show the importance of signs and pavement markings in increasing driver yielding.

In summary, the different intersection and treatment of traffic signals will influence both pedestrian and driver behavior, for example, the geometry will influence vehicle turning, thus creating different outcomes for a pedestrian crossing an intersection. Importantly, treatments such as marked cross walks and signs as well as flashing light beacons all contribute to an environment where drivers yield to pedestrians.

APPENDIX B: DATA COLLECTION

DATA COLLECTION UNITS

Video data were collected using the Minnesota Traffic Observatory (MTO) Traffic Information Monitors (TIMs). The TIMs are specially designed by the MTO for traffic data collection in the field and have been extensively tested in prior data collection deployments by the MTO.

Each TIM is equipped with a RaspberryPi microcomputer for data logging, a battery pack that allows for continuous operation for up to 14 days, a 30-foot extendable mast, and a video camera, and a water-tight enclosure. A TIM is shown in Figure below. Video data is recorded to USB memory sticks that can be retrieved to obtain the collected data.

TIMs are affixed with a large warning sign to deter members of the public from interfering with the data collection. The sign reads “Danger High Voltage”. However, this sign is purely for deterrence since the TIMs operate on 12-volt batteries and do not pose a risk to the general public. Recently, the TIMs have also been equipped with an additional sign that indicates that no personal information is recorded and that this data collection is not associated with law enforcement.



Figure B.1: TIM unit deployed showing water tight enclosure at bottom containing microcomputer and batteries as well as extendable mast and camera.

DATA COLLECTION PROTOCOL

TIMs were deployed for two weeks at a time to maximize the potential number of pedestrian-driver interactions observed at each deployment site. While not all video data may be used, all video data will be stored to be used in the future if necessary.

Each TIM must be attached to an existing mast or pole for stability. Encroachment permits are obtained from the relevant jurisdiction. TIMs are attached to the mast using padded bracing that does not damage the existing infrastructure while securing the TIM and ensuring it does not pose a tipping hazard.

Once the TIM is positioned, the camera is adjusted to capture the relevant crossing and approach, and the mast is raised to obtain the correct field of view.

DATA COLLECTION SITES

Note that all Figures in this appendix are oriented with the North at the top.

Site 1: 22nd Ave NE and University Ave, Minneapolis, MN

Date of Deployment: 9/21/2022 – 10/2/2022

Notes: The TIM was installed attached to a tree on the northwest corner of 22nd Ave NE and University Ave NE. This site included an offset of the minor road (22nd Ave NE) as it crosses through the intersection.



Figure B.2: Data collection site at 22nd Ave NE and University Avenue, Minneapolis.

Site 2: Penn Ave S and W 63rd St, Richfield, MN

Date of Deployment: 9/21/2022 – 10/2/2022

Notes: The TIM was installed attached to a light post in front of the Lund's and Byerly's grocery store northwest of the intersection. Note that this is a T intersection with a one-way driveway that exits from the parking lot.



Figure B.3: Data collection site at Penn Ave S and W 63rd St, Richfield, MN.

Site 3: Cleveland Ave and Pinehurst Ave, St. Paul, MN

Date of Deployment: 9/23/2022 – 10/2/2022

Notes: The TIM was installed by a light post on the southwest corner of the intersection. Due to the proximity of the TIM to the intersection, a wide-angle fisheye lens was used to capture the entire scene.

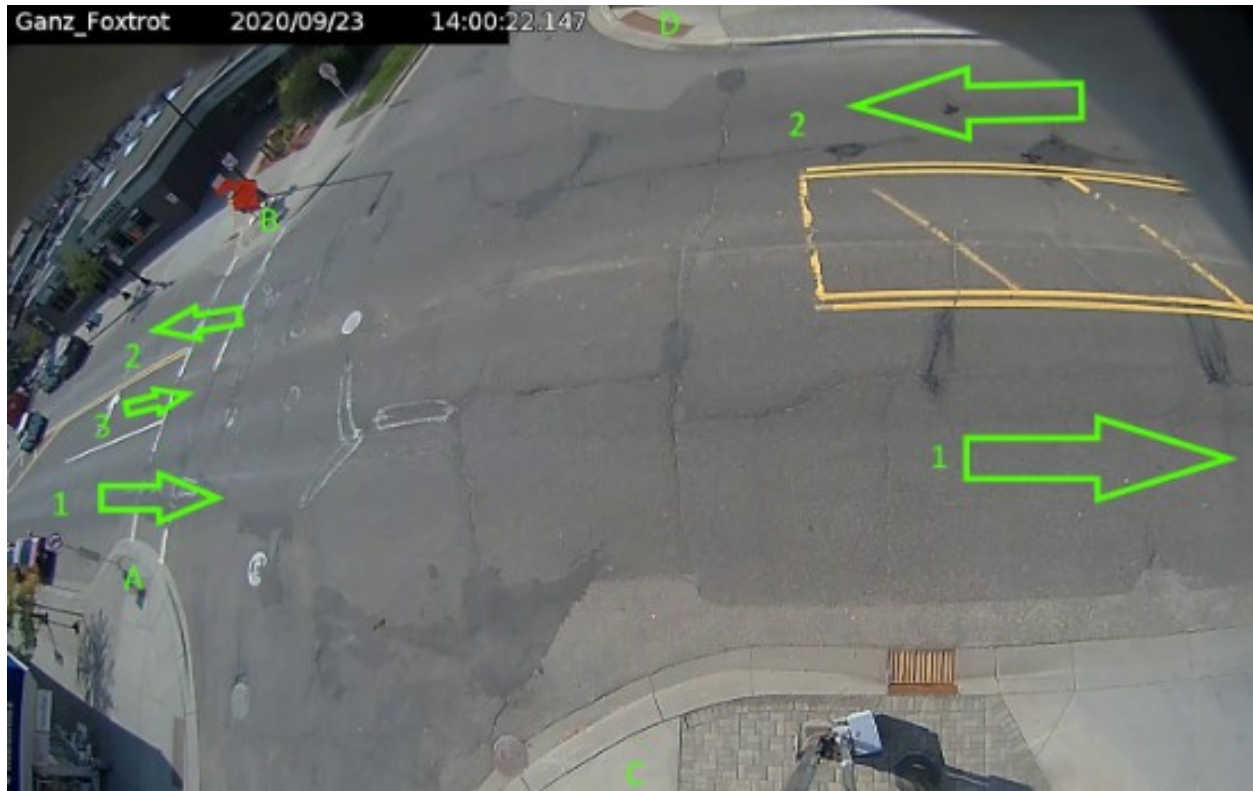


Figure B.4: Data collection site for Cleveland Ave S and W Pinehurst Ave, St. Paul, MN.

Site 4: W 66th St. and Grand Ave S, Richfield, MN

Date of Deployment: 5/28/2021 - 6/15/21

Notes: TIM was deployed on the light post shown as below with the intention of observing pedestrians crossing the intersection. This was set up on 5/28. This seemed to be the spot where pedestrians are most likely to cross given the presence of ADA pads and the median making it easier for pedestrians to cross.

Grand Ave and 66th St, Richfield, MN



Figure B.5: Data collection site at 66th and Grand, Richfield, MN.

Site 5: E 66th St. and Elliot Ave, Richfield, MN

Date of Deployment: 5/28/2021 - 6/15/2021

Notes: This location features two different cameras that are mounted on the same machine. The cameras face opposite directions with the purpose of recording the intersections of both E 66th St and Elliot Ave, as well as E 66th St and Chicago Ave. This was done because these two crossings have all of the same variables except for the fact that the crosswalk near Elliot Ave is painted and the one near Chicago Ave is not painted.

Elliot Ave and 66th St, Minneapolis, MN



Figure B.6: Data collection site at E 66th St. and Elliot Ave, Richfield, MN.

Site 6: E 66th St. and Chicago Ave, Richfield, MN

Date of Deployment: 5/28/2021 - 6/15/2021

Notes: As stated above, this data collection site used a single TIM with two cameras to collect data at two sites simultaneously. Importantly, the similarity between the sites allows for direct comparison between the marked and unmarked crosswalk.

Chicago Ave and 66th St, Richfield, MN



Figure B.7: Data collection site at E 66th St. and Chicago Ave, Richfield, MN.

Site 7: Portland Ave and E 74th St. Richfield, MN

Date of Deployment: 6/1/2021- 6/15/21

Notes: This camera was set up in a residential area near a marked crosswalk. This camera captured pedestrian and driver interactions mostly along the crosswalk that is outlined below but was also able to capture some interactions along the crosswalk on the far side as well.

Portland Ave and 74th St, Richfield, MN



Figure B.8: Data collection site at Portland Ave and E 74th St. Richfield, MN.

Site 8: County B Rd W and Pascal St. Roseville, MN

Date of Deployment: 6/3/2021- 6/11/2021

Notes: This camera was mounted on June 3rd to a road sign on the corner of the intersection. The set up was challenging given the close proximity of the camera and the walkway.

County Road B and Pascal St, Roseville, MN



Figure B.9: Data collection site for County B Rd W and Pascal St. Roseville, MN.

Site 9: Division St and 6th St. E, Northfield, MN

Date of Deployment: 6/15/2021 - 6/24/2021

Notes: This camera was mounted on June 15th to a light post to the southwest of the intersection. The camera was raised to about 20 feet to get a nice view of most of the intersection.

Division Ave and 6th St, Northfield, MN



Figure B.10: Data collection site at Division St and 6th St. E, Northfield, MN.

Site 10: Division St and 3rd St. E, Northfield, MN

Date of Deployment: 6/15/2021 - 6/30/2021

Notes: This camera was mounted on June 15th to a light post to the southwest of the intersection. The camera was raised to about 20 feet to get a nice view of most of the intersection.

Division Ave and 3rd St, Northfield, MN



Figure B.11: Data collection site at Division St and 3rd St. E, Northfield, MN.

Site 11: Nicollet and West 73rd St, Richfield, MN

Date of Deployment: 6/16/21 - 6/24/21

Notes: This camera was mounted on June 16th to a sign on the Northside of the intersection. This intersection was a bit tricky because there were no ideal mounting spots. The mounting spot that we settled with was on a sign right in the median of the road, just north of the intersection.

Nicollet Ave and 73rd St, Richfield, MN



Figure B.12: Data collection site at Nicollet and West 73rd St, Richfield, MN.

Site 12: E 38th St. and 24th Ave, Minneapolis, MN

Date of Deployment: 6/30/2021 - 7/11/21

Notes: This camera was mounted on June 30th to a light post to the northwest of the intersection.

24th Ave and 38th St, Minneapolis, MN



Figure B.13: E 38th St. and 24th Ave, Minneapolis, MN.

Site 13: Snelling Ave S and Stanford Ave, St. Paul, MN

Date of Deployment: 7/1/2021 - 7/15/21

Notes: This camera was mounted on July 1st to a light post to the northwest of the intersection. The camera was raised to about 20 feet to get a nice view of most of the intersection. The light post was covered by a tree but the camera was situated just below the leaves of the tree.

Stanford Ave and Snelling Ave, St. Paul, MN



Figure B.14: Data collection site at Snelling Ave S and Stanford Ave, St. Paul, MN.

Site 14: Dale St. and Sherburne Ave, St. Paul, MN

Date of Deployment: 7/1/2021 - 7/15/21

Notes: This camera was mounted on July 1st to a light post to the northwest of the intersection. The camera was raised to about 12 feet to get a nice view of most of the intersection. The light post was covered by a tree but the camera was situated just below the leaves of the tree.

Sherburne Ave and Dale St, St. Paul, MN

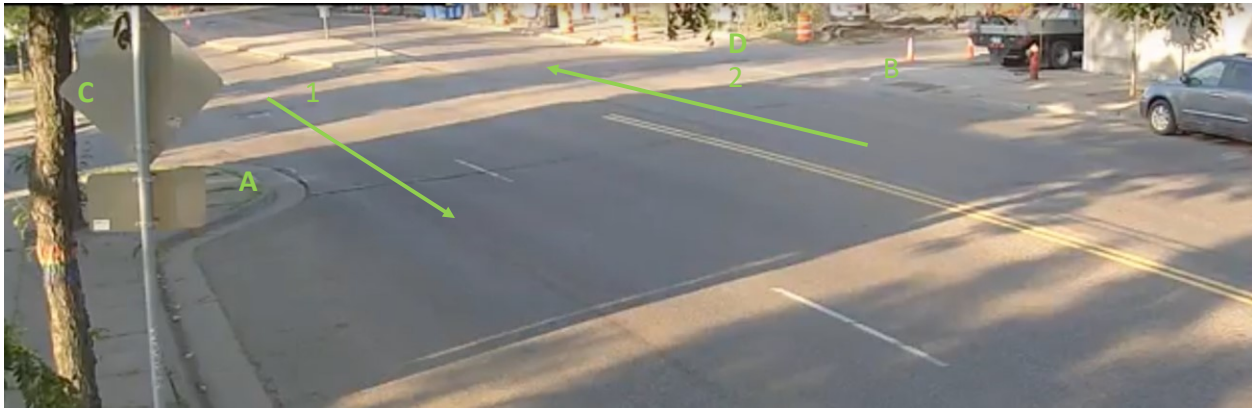


Figure B.15: Data collection site at Dale St. and Sherburne Ave, St. Paul, MN.

Site 15: N Upton Ave and Lowry Ave N, Minneapolis, MN

Date of Deployment: 7/15/2021 - 7/29/2021

Notes: This camera was mounted on July 15th to a light post to the northeast of the intersection. The camera was raised slightly but given its distance from the middle of the intersection, a wide angle lens was required to capture the full intersection.

Lowry Ave and Upton Ave, Minneapolis, MN



Figure B.16: Data collection site at N Upton Ave and Lowry Ave N, Minneapolis, MN.

Site 16: W 52nd St and France Ave S, Edina, MN

Date of Deployment: 7/15/2021 - 7/29/21

Notes: This camera was mounted on July 15th to a light post to the Southeast side of the intersection. The Camera was deployed onto a tree near the intersection and raised about 10 feet in the air for a good view of the intersection.

France Ave and 52nd St, Minneapolis, MN



Figure B.17: Data collection site at W 52nd St and France Ave S, Edina, MN.

Site 17: 9th Ave NE and Marshall St. NE, Minneapolis, MN

Date of Deployment: 7/15/2021 - 7/29/21

Notes: This camera was mounted on July 15th to a light post to the East side of the intersection. The camera was raised about 10 feet to get a good view of any interactions. This is a relatively busy intersection with high pedestrian volume.

Marshall Ave and 9th St, Minneapolis, MN



Figure B.18: Data collection site for 9th Ave NE and Marshall St. NE, Minneapolis, MN.

Site 18: Lowry Ave and James Ave Minneapolis, MN

Date of Deployment: 9/25/2021 - 10/5/2021

Notes: This site was originally instrumented for data collection with the TIM on July 15 to July 29. However, as a result of a data logging error, no data was recorded during this time. This site was re-instrumented on September 25th, 2021. The camera was placed on the northeast corner of the intersection and was left there for 11 days.

Lowry Ave and James St, Minneapolis, MN



Figure B.19: Data collection site at Lowry Ave and James Ave Minneapolis, MN.

APPENDIX C: IDENTIFIED FEATURES

The following interaction features are identified for each interaction:

- **Location_ID:** numeric ID (1-18) to differentiate locations for analysis.
- **location_name:** Intersection roads (e.g., County B. Road W and Pascal St.)
- **Date:** observation data in year, month, day format (e.g., 20210603 for June 3, 2021)
- **Time Showed Intent:** the time that the pedestrian first arrives at the intersections and indicates intent to cross presented in hour, minute, second format (e.g., 193825 for 7:38:25pm)
- **Time Started Crossing:** the time that the pedestrian begins the crossing presented in hour, minute, second format (e.g., 193825 for 7:38:25pm)
- **Time Finished Crossing:** the time that the pedestrian reaches the sidewalk on the other side of the intersection presented in hour, minute, second format (e.g., 193825 for 7:38:25pm)
- **Number of Pedestrians:** the number of pedestrians in the party
- **Yielder:** V for driver yields, P for pedestrian yields
- **Close Call:** 1 if the intersection is perceived as a close call, 0 if not
- **Pedestrian Origin:** letter A-D indicating which intersection corner the pedestrian originated at – see Deliverable 6 for full details
- **Pedestrian Destination:** letter A-D indicating which intersection corner the pedestrian ends at – see Deliverable 6 for full details
- **Pedestrian Stop Midway:** 1 if the pedestrian is forced to stop during the crossing to allow a vehicle to pass, 0 if the pedestrian is able to walk without having to yield to another vehicle mid-way
- **Interaction/Event Type:** A: pedestrian crossed at a comfortable pace as soon as a gap was present; B: pedestrian crossed at fast pace to utilize small gap as soon as a gap was present; C: pedestrian waited after gap was present before crossing at a comfortable pace; D: pedestrian waited after gap was present before crossing at fast pace; E: pedestrian did not wait for gap before crossing
- **Pedestrian type:** A: person on foot; B: person on bike; C: person on vehicle (e.g., scooter/hoverboard); D: person walking bike; E: mix of pedestrian types; F: other (see comments for entry); G: person with a dog; H: person with a stroller or small child
- **Vehicle Speed:** estimate of vehicle speed (mph)
- **Opposite Direction Yield:** 0 if no opposite direction vehicle, 1 if the vehicle traveling in the opposite direction yields, 2 if the vehicle traveling in the opposite direction does not yield
- **Following Vehicle:** 1 if there was a vehicle behind the interaction vehicle, 0 if not
- **Posted Speed:** Posted speed limit (mph)
- **Number of Lanes at Main Street:** number of lanes that the pedestrian is crossing
- **Crossing Width (Major):** pedestrian crossing distance (ft)
- **Bike Lane(s):** 0 if no bike lanes present on the major road, 1 if there are bike lanes present
- **Weather:** 0 if there is no precipitation; 1 if it is raining; 2 if it is snowing
- **Signage:** 0 if there are no signs for the crosswalk, 1 if the crosswalk is signed
- **Markings:** U for unmarked, C for continental (zebra) markings, S for standard (two solid white lines)

- **Presence of Single Family Housing:** adjacent land use indicating presence of single family homes within a 1 block radius of the intersection, 1 for present, 0 for not present
- **Presence of Multi-family Housing:** adjacent land use indicating presence of multifamily homes within a 1 block radius of the intersection, 1 for present, 0 for not present
- **Presence of Commercial:** adjacent land use indicating presence of commercial buildings within a 1 block radius of the intersection, 1 for present, 0 for not present
- **Presence of Gas Station/Convenient Store:** adjacent land use indicating presence of a gas station or convenience store within a 1 block radius of the intersection, 1 for present, 0 for not present
- **Presence of Restaurants/Bars:** adjacent land use indicating presence of bars or restaurants within a 1 block radius of the intersection, 1 for present, 0 for not present
- **Presence of Parking Lots:** adjacent land use indicating presence of large street-facing parking lots (not on-street parking) within a 1 block radius of the intersection, 1 for present, 0 for not present
- **Lighting:** 0 for no lighting, 1 for lighting
- **Road surface:** 0 for dry, 1 for wet
- **# of bus stops near the intersection:** number of bus stops within one block of the intersection on the main road
- **Minor AADT:** annual average daily traffic on the minor road
- **Major AADT:** annual average daily traffic on the major road
- **Dist. to Nearest Park:** distance in miles to nearest park
- **Dist. to Nearest School:** distance in miles to nearest school
- **Presence of on street parking:** presence of on-street parking, 0 – no parking; 1 – one-sided parking; 2 – two-sided parking
- **PAWS Score:** PAWS score, see <https://mndot.maps.arcgis.com/apps/View/index.html?appid=1cc55aa66d3844a98402c84673f73d14> for more details
- **Tree Cover:** 0-4 indicating the number of crossing points covered by trees
- **Comments:** a field for any additional comments as noted during data extraction

Additionally, the following are identified for each interaction for vehicles traveling in both for Direction 1 (vehicles traveling toward the camera) and Direction 2 (vehicles traveling away from the camera):

- **Is there a vehicle?:** Presence of vehicle in this direction, 1 for yes, 0 for no
- **Number of Vehicles Not Yielding:** total number of vehicles that do not yield in this direction of travel
- **Type of not yielding driver vehicle – vehicle 1:** type of the first not-yielding vehicle: 1 for sedan, 2 for SUV, 3 for utility vehicle (delivery/service vehicle), 4 for truck, 5 for bus
- **Type of not yielding driver vehicles:** list of the type of subsequent not-yielding vehicles: 1 for sedan, 2 for SUV, 3 for utility vehicle (delivery/service vehicle), 4 for truck, 5 for bus (e.g., 1, 1, 2 for sedan, sedan, bus)

- **Was there a yielding driver?:** 1 for if there was a yielding driver in this direction, 0 for if there was not a yielding driver in this direction
- **Type of yielding driver vehicle:** yielding driver vehicle: 1 for sedan, 2 for SUV, 3 for utility vehicle (delivery/service vehicle), 4 for truck, 5 for bus
- **Type of Yield:** S for a vehicle that comes to a complete stop to allow the pedestrian to cross, R for a rolling yield where the driver decreases speed but does not come to a complete stop, O for any other form of yielding (e.g., changing driving path to avoid a pedestrian), N or blank entry for a non-yielding driver
- **Following vehicle:** 1 if there is a vehicle behind the yielding driver, 0 if not

Note: cells that are left empty are either not applicable (e.g., no vehicle to report) or were not collected for that interaction technical reasons. Often, this is because a sight obstruction made it difficult to accurately assess.