Study of Pedestrians' Unsafe Road Crossing Behaviour at Signalised Intersection Crosswalks

Thesis submitted in partial fulfilment of the requirements
for the award of the degree of

Doctor of Philosophy

by

Rahul Raoniar

(Roll No. 166104036)

Under the Guidance of

Prof. Akhilesh Kumar Maurya

Department of Civil Engineering

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

Guwahati - 781039, India

November 9, 2022

Study of Pedestrians' Unsafe Road Crossing Behaviour at Signalised Intersection Crosswalks

Thesis submitted in partial fulfilment of the requirements
for the award of the degree of

Doctor of Philosophy

by

Rahul Raoniar

(Roll No. 166104036)

Under the Guidance of

Prof. Akhilesh Kumar Maurya

Department of Civil Engineering

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

Guwahati - 781039, India

November 9, 2022

Dedicated to My Family

Certificate

This is to certify that the thesis entitled **"Study of Pedestrians' Unsafe Road Crossing Behaviour at Signalised Intersection Crosswalks"**, submitted by **Rahul Raoniar** [Roll No. 166104036], to the Indian Institute of Technology Guwahati, for the award of the degree of **Doctor of Philosophy** in Transportation Systems Engineering, is a record of the original, bona fide research work carried out by him under my supervision and guidance. The thesis has reached the standards fulfilling the requirements of the regulations related to the award of the degree.

The results contained in this thesis have not been submitted in part or in full to any other University or Institute for the award of any degree or diploma to the best of our knowledge.

21-11-2022

Prof. Akhilesh Kumar Maurya
Department of Civil Engineering
Indian Institute of Technology Guwahati

Declaration

I declare that this written submission represents my ideas in my own words, and where others' ideas and words have been included, I have adequately cited and referenced the original source. I declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated, or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will cause disciplinary action by the Institute and can also evoke penal action from the source which has thus not been properly cited or from whom proper permission has not been taken when needed.

Rahul Raoniar

. . . .

Rahul Raoniar

Roll No.: 166104036

Date: 09-11-2022

Place: Guwahati

Acknowledgements

I am thankful to my Supervisor **Prof. Akhilesh Kumar Maurya** for his continuous support, conviction, encouragement and invaluable advice in the course of the completion of my dissertation. I have been blessed to have an understanding guide who always bestowed freedom to explore ideas. He is one of the best person I have come across in my life, who tough me how to survive in tough times and enjoy the little achievements.

I am sincerely grateful to **Prof. Rajan Choudhary** for his insightful comments at different stages of my research work. I would like to thank doctoral committee members for their constant support and encouragement.

The influence of those with whom I worked and interacted during my journey at IIT Guwahati is also continues to be important. I would like to thank my friend, **Anant Aishwarya Dubey** and **Durgesh Kumar**, who continuously motivated me during my research work.

I would like to thank **Houyem Hafdi** for proofreading and the **anonymous reviewers** who reviewed my various publications drawn from this thesis. Their constructive comments have greatly improved my research work and underlying thesis.

The data collection and extraction for research would have been immensely difficult without the effort of **Upen Gohain** (Jr. Technical Superintendent, IIT Guwahati), **Sajad Maqbool** (Master's student, IIT Guwahati), **Abhiney Pathak** (Master's student, IIT Guwahati), **Mohit Chugh** (Master's student, IIT Guwahati) and **Sruthi Singh** (Master's student, IIT Guwahati) for which I am very much obliged. I would like to thank my father **Shibsankar Raoniar**, who helped me obtain permission for data collection from Kolkata Traffic Police and ushered me to conduct a preliminary survey and survey site selection across Kolkata city.

I would also like to thank Kolkata Traffic Police for their support and cooperation, especially for providing permission to conduct video graphics and questionnaire surveys.

Finally, my deepest thank goes to my parents (**Shibsankar Raoniar** [father] and **Sova Raoniar** [mother]), my brother (**Subham Raoniar**) and my Grand Mother (**Urmila devi**) for constantly motivating me in the course of my higher education.

Rahul Raoniar

Abstract

The movement of pedestrians in an urban environment is vital for promoting safe and comfortable sustainable transport witing the existing transportation system. To enable a safe and comfortable walking experience, various at-grade facilities were built, like crosswalks at mid-blocks and intersections with proper marking and signage. The inadequacy of intersection crosswalks in terms of design and dimensions increases the woes of pedestrians. These impediments render the performance of these facilities unsatisfactory to most pedestrians, leading to risk-taking behaviour and subsequent exposure to danger. Thus it is very clear that understanding the pedestrian need by investigating the existing road crossing behaviour is prime important for planners and policymakers to enable a safe, comfortable and convenient pedestrian environment.

It is essential to understand the past research in pedestrian road crossing behavioural studies in general and especially at signalised intersection crosswalks. Therefore, the overview of existing pedestrian behaviour in the road environment was explored by conducting a detailed review of pedestrian risk-taking behaviour. Past studies mostly explored various non-social level attributes, and few tried to understand the influence of social level attributes on risk-taking behaviour. To overcome this research gap, the current study tried to understand what motivates/forces pedestrians to take risks at signalised intersection crosswalks in the light of social and non-social attributes. Real-world observational data were gathered from three one-way crosswalks in Kolkata city to investigate various hypotheses. The study results highlighted that, similar to non-social attributes, the social level attributes are also crucial for understanding pedestrian risk-taking behaviour.

In the 21^{st} century, a digital distraction like mobile use in road environment has increased the number of injuries among the young population. There was very limited research available in developing countries like India that tried to understand the behavioural differences induced due to various types of distractors like mobile talking, texting, headphones use, group talking and eating/drinking/smoking while crossing the road. To fill this research gap, the current study investigated the influence of different types of distraction on road crossing behaviour at three signalised intersection crosswalks. Binary logistic regression has been utilised to understand

the behavioural difference between pedestrians with and without distraction. The study results highlighted that various distractors had different influences on road crossing behaviour regarding crossing speed, signal violation, conflict with opposing pedestrians and traffic glance.

Longer red-phase lengths led to the long waiting time at intersection crosswalks. The longer waiting duration often does not satisfy the threshold waiting time of individual pedestrians, which leads to signal non-compliance. Understanding the optimal waiting duration is essential for designing the pedestrian signal phase. Very limited research is available that tries to understand the pedestrian waiting behaviour at intersections using duration data. Thus, the present study investigated the waiting duration of pedestrians using the time-to-event approach (survival analysis) using data collected from eight two-way signalised intersection crosswalks in Kolkata. From the study results, it was observed that the Log-normal (Log-normal Accelerated Failure Time model) distribution best represented the waiting duration of pedestrians at intersection crosswalks in Kolkata city. Further, a high danger rate is shown by short waiting time (less than 3 s), while endurance towards waiting was reflected in the span of 36.7 s. Further, pedestrian glance was the most crucial predictor of waiting behaviour.

Conflict or interaction-based studies are preferred for pedestrian safety evaluation due to the non-availability of historical crash data in developing countries. Thus, the present study evaluated pedestrian crossing safety using the interaction-based technique to obtain the Pedestrian Safety Margin (PSM). Two regression models were developed for one-way (multiple linear regression model) and two-way (random intercept regression model) crosswalks. From the PSM model, it was observed that crossing speed, type and size of the vehicle, approaching direction, location of interaction and number of traffic plying on the street affected the PSM.

The study results will be helpful for traffic engineers, planners and enforcement agencies to formulate new policy guidelines to promote safe and comfortable pedestrian infrastructure. It will also help in assessing and prioritising the requirements of pedestrians at existing intersections.

Contents

List of Tables

List of Figures

Abbreviations

AFT	**A**ccelerated **F**ailure **T**ime
AI	**A**rtificial **I**ntelegence
AIC	**A**kaike **I**nformation **C**riteria
AME	**A**verage **M**arginal **E**ffects
ANN	**A**rtificial **N**eural **N**etwork
AOV	**A**nalysis **O**f **V**ariance
APS	**A**ccessible **P**edestrian **S**ignal
BBG	**B**epin **B**ehari **G**anguly
BIC	**B**ayesian **I**nformation **C**riteria
CBD	**C**entral **B**usiness **D**istrict
CI	**C**oncordance **I**ndex
CPH	**C**ox **P**roportional **H**azard
DOF	**D**egree **O**f **F**reedom
EDA	**E**xploreatory **D**ata **A**nalysis
E/D/S	**E**ating/ **D**rinking/ **S**moking
GPO	**G**eneral **P**ost **O**ffice
HR	**H**azard **R**atio
KM	**K**aplan **M**eier
LL	**L**og **L**ikelihood
LED	**L**ight **E**mmiting **D**iode
LMICs	**L**ower and **M**iddle **I**ncome **C**ountires
LRT	**L**ikelihood **R**atio **T**est
ML	**M**achine **L**earning
MDS	**M**illion **D**eath **S**tudy
MLR	**M**ultiple **L**inear **R**egression
PSM	**P**edestrian **S**afety **M**argin

RHIME	**R**outine, **R**eliable, **R**epresentative, **R**esample **H**ousehold **I**nvestigation of **M**ortality with **M**edical **E**valuation
RTI	**R**oad **T**raffic **I**njury
SD	**S**tandard **D**eviation
SP	**S**mart **P**hone
SW	**S**hapiro-**Wilk**
USA	**U**nited **S**tates of **A**merica
VIF	**V**ariance **I**nflation **F**actor
VMS	**V**ariable **M**essage **S**ign

Symbols

H_i	The i_{th} Hypothesis
V	Video Graphic survey
Q	Questionnaire survey
O	Observational survey
S	Required sample size
N	Population size
e	Margin of error
z	Z-score
p	Percenrage occurance of a state
χ^2	Chi-square
r	Pearson's Correlation
$U(X\|Y)$	Theil's U (association) between X and Y
$S(X)$	Entropy
$S(X\|Y)$	Conditional entropy of X given Y
ρ^2	Pseudo rho-square
β	Model coefficient
β_k	The estimated value of k^{th} parameter
X_k	The k^{th} explanatory variable
$LL(\beta_0)$	Log-likelihood of a null model when all parameters set to zero
$LL(\hat{\beta})$	Log-likelihood of the model with all fitted parameters on convergence
X_L	The Likelihood Ratio Test
$\ln L$	Maximised log-likelihood of the model
k	Number of parameters estimated
s	Time in seconds
P_i	Probability of an event occuring
T	Pedestrian waiting time in seconds

t	Time till the event
f_u	Probability density function
$S(t)$	Survival function
t_i	Time when at least one signal violation event occurs
d_i	The number of signal violations events that took place in time t_i
n_i	the pedestrians who waited (survived) up to time t_i
p_i	Survival probability
$h(t, X)$	Risk rate at a particular instance
$h_0(t)$	Refernce risk function
λ	Acceleration Factor
ϵ_i	The residual in the log-transformed survival times or prediction error
μ	Intercept
σ	Scale parameter
S_{ei}	The survival function of normal distribution
$\phi(\epsilon)$	The distribution function for normal distribution
$H_{\epsilon i}(\epsilon)$	The cumulative hazard function
e^{β}	Odds/Hazard ratio
Y	The observed value of PSM
$\hat{Y}$	The predicted value of PSM
$\bar{Y}$	The Mean value of PSM
F	F-statistics
SS_R	Regression sum of squares
SS_E	Residual sum of squares
SS_T	Total sum of squares
q	Degree of freedom for regression analysis
R^2	R-squared value
R_i^2	The unadjusted multiple regression coefficient of the i_{th} variable
m/s	Meter per second
u_j	Group specific random intercept
ϵ_{ij}	Individual error term
ρ	Variance Partitioning Coefficient
W	Shapiro-Wilk test statistics
z	Standard normal score for estimate

Chapter 1

Introduction

1.1 Overview

Walking is the most effective and environment-friendly mode of transport for shorter trips. Walking forms an integral part of urban transport. In India, approximately 1-2 km length of urban trips is performed on foot daily (IRC:103, 2012). Due to the extensive demands of vehicular traffic, most of the researchers investigated the operational condition of vehicular traffic and only very few of them have addressed the issues related to non-motorised mode. In urban areas, significant attention is being given to walking due to rising levels of congestion and pollution.

Walking is the fundamental mode used for commuting to shorter trips which involve road crossing, which is a hazardous activity due to the interaction involved with motorised traffic. To reduce this interactivity, sidewalks, crosswalks, intersections, foot over bridges and skywalks are provided to pedestrians to encourage a safe walking environment. Among all, intersections are considered the most important elements of road transport networks. Though intersections were built to improve the safety, convenience, travel speed, capacity and efficiency of the vehicular road network. They are also an integral part of non-motorised transport, especially for pedestrians, as they provide a designated place for safe crossing.

The intersections are far more complex and involve high-risk elements for pedestrians, as pedestrians and vehicles share the common road space for movement. The intersections are divided into two categories, controlled and uncontrolled. It has been observed that urban signalised intersections have high pedestrian and traffic movement (e.g., Kolkata). However, a very minimal number of safety measures are being allocated for pedestrian safety. Various types of control strategies, such as fixed-time control signals, coordinated signals, dynamic control signals and adaptive controls, are allocated based on traffic and pedestrian flow characteristics. Fixed time signals are very common in Indian cities, while adaptive signals are rising in metropolitan cities (e.g., Kolkata). Even though the signal control mechanism provides right-of-way to pedestrians for safe crossing during the pedestrian green phase, still past studies highlighted that the road crossing risk involved is more severe in a signal-controlled intersection than in uncontrolled intersections. This is because drivers tend to break traffic rules in lower-middle-income countries. In addition, the allowance of turning vehicles during the pedestrian green walk phase increases the chances of conflicts. These circumstances might have contributed to a large number of pedestrian-vehicle crashes in developing countries like India (Kumar et al., 2019; Jensupakarn and Kanitpong, 2018; Yang and Najm, 2007; Koepsell, 2002).

In Lower-Middle-Income Countries (LMICs) like India, to fend off the interaction of traffic and pedestrians, foot-over bridges and underpasses are being constructed in pedestrian-concentrated areas to ensure safe movements while maintaining a smooth flow of vehicular traffic. At first glance, this seems to have resolved the pedestrian safety problem. However, studies showed that the desired effect has not been met yet. Inappropriate standards, poor maintenance, presence of beggars and obscured safety due to shielding by advertisements were cited as the main reasons (Nukta Ramadani et al., 2018). Dass et al. (2015) reported that the absence of proper lighting at night, narrow stairs, absence of ramps and lifts, and absence of a rest staircase discouraged pedestrians from using the overpass. Rankavat and Tiwari (2016) stated that despite pedestrians having unsafe perceptions of safety while crossing at grade (at crosswalks), they still prefer to use crosswalks more than over or underpass.

As most pedestrians prefer or are obliged to use crosswalks instead of foot over-bridges, the inadequacy of crosswalks in terms of design or dimensions has reduced pedestrian safety. The extended red phase and/or too-short green phase for pedestrians caused excessive delays for people. It appeared to be a key issue behind the risk-taking behaviour (signal violation). Additionally, bad surface conditions, improper crosswalk markings, poor association between crosswalks and sidewalks, inappropriate provision of guardrails and refuge islands, and allowance of free left turns at intersections contributed to inconvenience for pedestrians (Bansal et al., 2018). The other peculiar challenges pedestrians face are on-street parking at or near pedestrian facilities, and poor police enforcement adds to the users' woes (Mukherjee and Mitra, 2020). These impediments render the performance of these facilities unsatisfactory to most pedestrians, leading to risk-taking behaviour and subsequent exposure to danger. A recent study conducted in Kolkata city confirmed that an association does exist between pedestrian risk-taking behaviour (signal violation) and fatal road crash frequency (Mukherjee and Mitra, 2020).

Safety concerns for pedestrians are severe in some Indian cities, where pedestrians represent half of all fatalities (Mohan et al., 2020). The Indian official estimates revealed that pedestrian fatalities constitute 15% of total Road Traffic Injuries (RTI) (MoRTH, 2018), which is lower than other independent researchers' estimates. Hsiao et al. (2013) studied MDS (Million Death Study) data of 1.1 million Indian households; these data were collected using a face-to-face interview (with a relative or close acquaintance of the deceased). In this study, based on an enhanced version of the verbal autopsy technique (known as the routine, reliable, representative, resample household investigation of mortality with medical evaluation or RHIME), about 37% of pedestrian RTIs were reported. Based on the time trends and variations among the Indian states for deaths due to road injuries, Dandona et al. (2020) reported 76,729 (35.1%) pedestrian deaths of all road traffic deaths in 2017.

As per "The Road Accident Report 2018", West Bengal (an Indian State) topped the list of pedestrian RTI fatalities in 2018 with 2,618 deaths, followed by Maharashtra with 2,515 reported deaths. The number of pedestrians killed on Indian roads has risen by 84% between 2014 and 2018 (MoRTH, 2018). Current statistics highlighted that 62 pedestrians die daily on Indian roads from road traffic-related injuries (Das, 2019). Further, past accidental investigations have reported that a significant portion of vehicle-pedestrian collisions occurred at intersection crosswalks. In Charlotte, North Carolina (USA), many crashes happen involving pedestrians at intersections near Light Rail Transit (LRT) stations (Pulugurtha and Srirangam, 2021). In China, for example, more than 50% of pedestrian crashes occur at signalised intersections (Chen et al., 2017; Ren et al., 2011), while in Montreal, Canada, vehicle-pedestrian collisions account for more than 60% of reported accidents (Brosseau et al., 2013). In a developing country like India, accidents at the intersection have reached beyond 60% of all pedestrian-related accidents (Bansal et al., 2018; Priyadarshini and Mitra, 2018) and pedestrian risk-taking behaviour, especially signal violation, has been cited as one of the significant reasons (Mukherjee and Mitra, 2020). This clearly highlights the need for more pedestrian safety studies to determine the reasons behind pedestrians' unsafe behaviour and the factors influencing their safety so that preventive steps can be taken to reduce fatal pedestrian-vehicle collisions.

1.2 Need of the Study

Enhancing pedestrian safety is one of the challenges faced by transportation planners and traffic engineers worldwide. Most of the complexity arrives due to the variation and uncertainty involved in the pedestrian road crossing behaviour. The complexity and diverse nature of pedestrian crossing behaviour and driver's behaviour are described in the following paragraphs.

There are three types of signal phases for pedestrians at signalised intersections: green phase, flashing green phase and red/do-not walk phase. During the green phase, pedestrians are allowed to cross the street. During the flashing green phase, pedestrians are not allowed to enter the crosswalk, but those who have already entered the crosswalk are allowed to complete their crossing. Even though during the green phase, pedestrian holds the right of way, they still face safety threats due to drivers' non-compliance with the stop signal. Further, the lack of enforcement in developing countries like India increases pedestrians' woes. For example, during the pedestrian green phase, vehicles often occupy pedestrian crosswalks which force pedestrians to walk outside the zebra crossing, which elevates the crossing risk.

In addition, the crossing risk increases further when pedestrians themselves decide to cross the road illegally, taking risks (i.e., during the red/do-not walk phase). Pedestrians decide to take the risk for two specific reasons: i) when there are very few vehicles plying on the street and ii) when the waiting duration for safe crossing (green phase crossing) exceeds pedestrian threshold waiting time due to longer red-phase length. In such circumstances, pedestrians use their judgement and increase their crossing speed to take safe gaps between vehicles to complete their manoeuvre.

Moreover, turning vehicles make crossing difficulty much worse and unsafe for pedestrians. Further, unsafe behaviour displayed by pedestrians during road crossing in the form of mobile phone distraction has raised concern for safety researchers.

In addition, inconsistent infrastructure across the transport networks forces pedestrians to take risks during road crossing. For example, broken or non-functional signal heads, unoptimised signal phases, unavailability of crosswalk markings and refuse island also forces pedestrian to make risky crossing choices. It is important to improve the pedestrian infrastructure quality and implement safety interventions at intersection crosswalks. Hence, there is a need to study pedestrian crossing and risk-taking behaviour at signalised intersection crosswalks under mixed traffic conditions.

1.3 Problem Statement

In the previous section, the existing pedestrian crossing behaviour and its complexity in Indian traffic conditions (which may lead to unsafe road crossing) are discussed. Analysing pedestrian crossing behaviour could be challenging due to the complexity and uncertainty involved in road crossing behaviour. Thus, there is a need to understand pedestrian crossing behaviour, especially the unsafe one in crosswalks at signalised intersections, using real-world empirical data collected from Indian cities.

Road crossing is a daily activity for pedestrians; however, it is considered a threat if caution is not practised. At uncontrolled intersections, pedestrians are often cautious to avoid injury while crossing and take marginally safe gaps. Nevertheless, at a controlled intersection, if the pedestrian signal is red, the chances of safe crossing opportunities are very limited; some pedestrians may be willing to cross, whereas others may wait for the green light. Past studies extensively studied signal violation behaviour using various non-social factors (such as traffic and road infrastructure). But the risk-taking behaviour is far more complex as other factors may also motivate pedestrians to make such risky decisions. Other key factors are social information (like the number of pedestrians waiting in a group or attempting to cross during the red-signal phase) and pedestrian characteristics (like gender, age, and walking speed) which need further exploration to understand their influence on risk-taking motive.

As vehicle drivers and pedestrians use a shared space for movement (especially at intersections); thus, similar to driver distraction (Horrey and Wickens, 2006; Pekker et al., 2011), the pedestrian could also engage in distraction, which reduces crossing safety. Past studies have displayed the potential impact of distracted road crossing behaviour on crossing safety risks. However, their conclusions were not validated with accident statistics due to the paucity of such data availability/inventory. Ralph and Girardeau (2020) raised similar concerns and pointed out that despite the lack of field evidence in terms of accidents data, scholars tend to describe distracted walking as a severe problem, which might misplace the key focus from the overall effort of improving safety for all road users. Therefore, in developing counties like India in the absence of distraction-related accident records, a questionnaire survey of a representative sample could provide an overall accurate initial estimate of pedestrians' distraction rate and its severity in daily commutes (while in a road environment).

Thus, it is also worth noting that in developing countries where road traffic rules and regulations are lenient, it is unrealistic to expect that every road user will comply with traffic rules and regulations and abstain from distraction. Thus, it is essential to develop and deploy new strategies and countermeasures that minimise distraction rather than attempt to eliminate such behaviour altogether. One of the best strategies is to understand and assess the severity of such incidents in a naturalistic setting using an observational study framework. Thus, parallel to collecting subjective data (questionnaire survey), it will also be worthwhile to explore the manner in which pedestrian distractive engagement could be attributed to real-world road crossing behaviour in a developing country, India in particular.

The risk-taking behaviour of pedestrians at intersection crosswalks is highly associated with waiting delays on arrival. The unoptimised signal phase, especially the longer red-phase length, is mainly responsible for the extended waiting time. Identifying the optimal red-phase length is a challenging task, as the intersections greatly vary based on the traffic and signal configuration. Further, the dynamic nature of signals increases the complexity of the analysis. Thus, traditional models often fail to capture the influence of duration on unsafe events. Hence, further understanding is required to understand the waiting duration of pedestrians in signal violation so that optimal red-phase duration could be proposed to enhance signal compliance at signalised intersection crosswalks.

Pedestrian unsafe risk-taking behaviour often leads to unsafe interaction with approaching vehicles. Modelling and quantifying such interaction is challenging, as an enormous set of variables contribute to pedestrian-vehicle interactions. Therefore, it is necessary to develop a robust pedestrian-vehicle interaction model using a broader set of empirical data collected from many signalised intersection crosswalks.

1.4 Study Objectives

The main aim of the study is to explore and analyse pedestrian road crossing behaviour at signalised intersection crosswalks. To achieve this aim, the following four research objectives are formulated:

Research Objective I: Understanding the impact of social and non-social information on pedestrian signal violation behaviour at signalised intersection crosswalks.

Under this, the following three hypotheses were investigated in the context of Low- and Middle-Income Countries:

(i) Hypothesis 1 (H_1): the crossing decision of an oncoming pedestrian depends on whether the pedestrian is waiting alone at the curb or waiting with others (other pedestrians are already waiting before the candidate pedestrian arrived or joined later) standing nearby.

(ii) Hypothesis 2 (H_2): the likelihood of crossing in the red-light phase increases when one finds others (a significant number of pedestrians) crossing successfully in the red-light phase from the same or opposite direction.

(iii) Hypothesis 3 (H_3): as the waiting time for safe crossing (time until the green-signal phase start) increases, the likelihood of signal violation also increases.

Research Objective II: Understanding the impact of different types of distractions in road crossing behaviour at signalised intersection crosswalks.

In this study, the following sub-objectives were investigated in the context of Low- and Middle-Income Countries:

(i) Understanding the purpose and motivation behind mobile phone use by pedestrians while crossing the road using a representative questionnaire survey.

(ii) Assessing the rate at which pedestrians were subjected to a near-miss or an accident during distracted walking/crossing using a representative questionnaire survey.

(iii) Identifying the rate of different types of distraction in an observational study setting using a video graphic survey.

(iv) Identifying the differences in road crossing behaviour (demographic, violation, exposure, state of crossing and cautionary/glance) between pedestrians with digital distraction (using mobile) and without any distraction.

(v) Identifying the impact of other types of distraction (such as pedestrians' social distraction when conversing in a group or consuming beverages/smoking, as they cross the street) on road crossing behaviour.

Research Objective III: Modelling pedestrian waiting behaviour at signalised intersections using time-to-event analysis (survival analysis).

The following sub-objectives were investigated under the study in the context of Low- and Middle-Income Countries:

(i) To understand the pedestrian waiting behaviour at signalised intersection crosswalks using a non-parametric Kaplan-Meier curve.

(ii) Understanding the impact of various covariates (especially traffic glance and distraction) using a semi-parametric Cox Proportional Hazard regression and parametric Accelerated Failure Time (AFT) model alongside the common covariates used in previous studies.

Research Objective IV: Modelling pedestrian-vehicle interaction at signalised intersection crosswalks using safety margin approach.

In the current study, the following sub-objectives were investigated in the context of Low- and Middle-Income Countries:

(i) To understand the impact of pedestrian demographic and crossing characteristics on Pedestrian Safety Margin (PSM).

(ii) To understand the impact of cautionary/glance and distraction behaviour on Pedestrian Safety Margin (PSM).

(iii) Understanding the impact of vehicle characteristics on Pedestrian Safety Margin (PSM).

1.5 Scope of the Study

The scope of the current research is limited to the understanding of pedestrians' risky road crossing behaviour at marked signalised intersection crosswalks in Kolkata city. The video graphic data collection was done only for off-peak hours (11 am — 2 pm) to avoid any interference from traffic police. In addition, the data were extracted for one-way crossing, i.e., towards the camera.

1.6 Thesis Organisation

This thesis consists of 9 chapters, including the present chapter (Chapter 1) of introduction to the research topic. This chapter describes the need for the study, the problem statement and objectives of the study and a brief outline of the thesis. Chapter 2 provides a review of past literature related to pedestrian signal violation, distracted road crossing, waiting behaviour at signalised intersection crosswalks and safety margin. Chapter 3 presents a brief description of the methodology adopted to fulfil the current objectives. Chapter 4 describes details of survey locations, data collection, data extraction and questionnaire survey. The pedestrian signal violation behaviour model is discussed in Chapter 5. Chapter 6 deals with distracted road crossing models. Chapter 7 describes pedestrian waiting behaviour modelling. Chapter 8 presents pedestrian-vehicle interaction modelling based on the safety margin approach. Chapter 9 summarises the conclusions and research contributions of the current work and also presents practical implications for improving pedestrian crossing safety at signalised intersection crosswalks.

Chapter 2

Literature Review

2.1 Introduction

To encourage a safe walking environment in the transportation network, there is a need for a clear understanding of pedestrian road crossing behaviour under various conditions. Understanding the pedestrians' real requirements for safe road crossings will help policymakers and planners develop better and more innovative solutions that may enhance pedestrian safety. Many studies focused on pedestrian road crossing behaviour at various types of facilities were conducted in developed and developing countries. However, limited studies analysed pedestrian crossing safety at intersection crosswalks in developing countries, such as India. Hence, it is important to study pedestrian road crossing behaviour using observational and empirical data to deploy up-to-date interventions.

Pedestrian road crossing behaviour significantly differs in developing countries (like India) as compared to developed countries due to shared mixed traffic space, outdated traffic rules and regulations, and lack of enforcement. This chapter presents a detailed review of all the studies which were conducted on pedestrian road crossing behaviour at signalised intersection crosswalks under the various heading presented in Figure 2.1. The chapter ends by summarising the outcomes of the literature review.

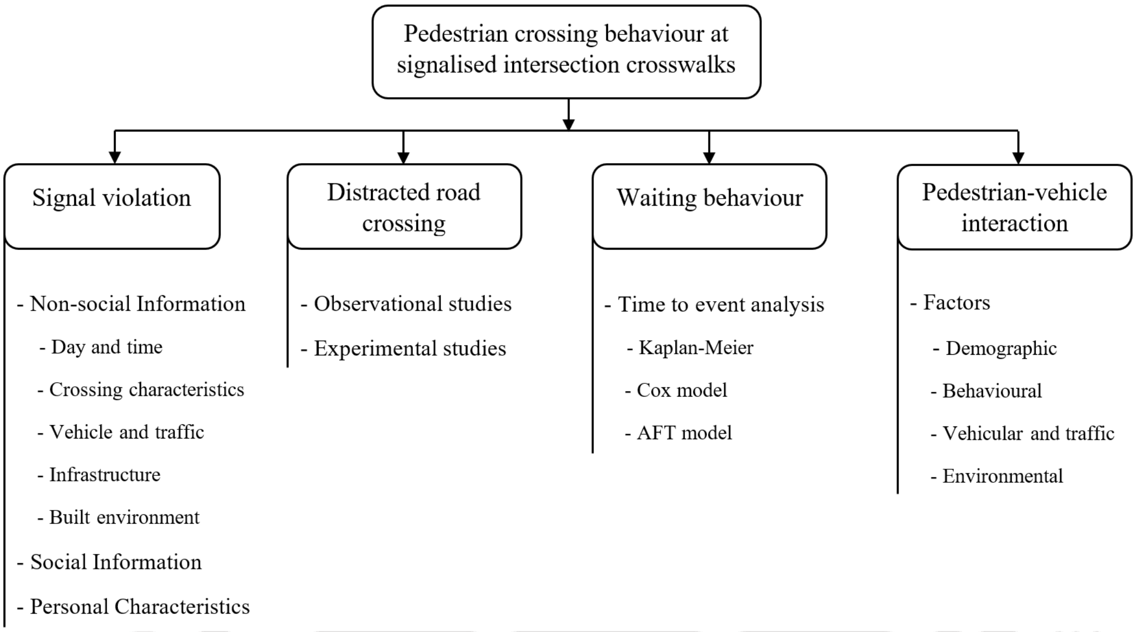


FIGURE 2.1: Literature review flow chart.

2.2 Pedestrian Signal Violation Behaviour at Signalised Intersections

Researchers have identified several key features (variables) that contribute to pedestrian risk-taking in previous studies. Non-social features (such as traffic and road infrastructure) and social features (like the number of pedestrians in a group attempting to cross during the red-signal phase) were among the potential features which significantly contribute to pedestrian road crossing decisions (Rosenbloom, 2009; Zhou et al., 2009). Apart from non-social and social features, pedestrian characteristics (like gender, age, and walking speed) were also observed to influence risk-taking behaviour. The influence of all three features (non-social, social, and pedestrian characteristics) has been discussed in detail in the subsequent paragraphs.

2.2.1 Non-Social Information

Several past studies have explored how non-social factors influenced pedestrians' road-crossing decisions. The non-social information source includes 'day and time of the week', 'pedestrian crossing' characteristics, 'vehicle and traffic' characteristics (speed, flow, and density), 'infrastructures', and 'built environment', i.e., road width, type of street, availability of marking, and signalised or unsignalised roads (Faria et al., 2010; Zhou et al., 2009).

(a) Day and Time of the Week: Previous studies revealed that the day of the week, the period of the year, or the time of day impact the proportion of violations (Wang et al., 2011), and they are essential predictors of a road collision. Based on the Indian epidemiological data, Kumar and Toshniwal (2015) reported that pedestrian accidents at night were more critical than daytime.

(b) Pedestrian Crossing Characteristics: Aside from personal characteristics, the amount of time a pedestrian spends waiting at the red light signal has been found to have a significant impact on their road crossing behaviour. The threshold waiting time for pedestrians differs according to different parameters, such as urgency for work (Guo et al., 2011), saving travel time (Forsythe and Berger, 1973; Ren et al., 2011), age (Brosseau et al., 2013), red-light phase length (Brosseau et al., 2013; Van Houten et al., 2007), availability of exclusive pedestrian signal, and whether crossing as an individual or in a group/platoon (Rosenbloom, 2009). Past studies highlighted that males were the most significant risk-takers in a road environment and had less patience upon waiting compared to female pedestrians (Rosenbloom, 2009). Tiwari et al. (2007) and Guo et al. (2011) used survival analysis and identified that pedestrian crossing activity and the risk of traffic violation were time-dependent. The results showed that the longer the time has passed since the waiting period began, the more likely pedestrians are to end their wait quickly in the red-signal phase. Two critical time points, 3 and 50 seconds were observed. The high danger rate shown by the short waiting time (less than 3 seconds) suggests that certain pedestrians crossed the street without waiting for a green light. When they arrived at the crosswalk, these pedestrians crossed the street immediately if the traffic lanes were clear of approaching vehicles. The pedestrian endurance towards waiting was reflected in the span of 50 seconds. Dommes et al. (2015) observed that pedestrians who were reluctant to wait (cross immediately after arrival) were more likely to look at traffic than signal as their signal violation decision was predetermined, and they needed to look for an adequate safe gap between vehicles. Further, few studies have quantified the impact of waiting time on signal violations, which highlighted that with an increase of 10% of the waiting time, the probability of violations and dangerous violations had increased by 7.9% and 2.1%, respectively (Brosseau et al., 2013; Chen et al., 2017).

(c) Vehicle and Traffic Characteristics: The pedestrian crossing also depends on the number of vehicles plying on the road. High traffic rate, according to Dommes et al. (2015), has mixed effects. On one side, the pedestrian green light may take longer to activate, which might cause pedestrians to breach the signal. While on the other side, crossing in heavy traffic is strenuous, which can minimise the number of red-light violations. Also, it has been found that the presence of parked vehicles in the immediate vicinity inhibited looking activities and increased the likelihood of crossing against the red light (Mukherjee and Mitra, 2020). Among the operational traffic attributes, vehicle volume (Harrell, 1991; Sun et al., 2004), available gaps (Xu et al., 2013; Koh and Wong, 2014) and vehicle speed (Hamidun et al., 2013; Retting et al., 2002) are observed to have a significant impact on pedestrian conformity to the signal timings.

(d) Infrastructure: Past studies highlighted that crosswalk marking, number of traffic lanes, the length of pedestrian crossings and type of pedestrian signal head impact pedestrians' red-light running rate. The impact of crosswalk markings on pedestrian signal violation was mixed. Koepsell (2002) studied the marked crosswalk in six U.S. cities and reported that markings had minimal effect on older pedestrians. Mukherjee and Mitra (2020) studied 55 potential signalised intersections in Kolkata (India) and observed that the absence of a marked pedestrian crosswalk was associated with a high share of illegal crossing. A significant number of studies pointed out that the crosswalk length was also a significant factor in pedestrians' violation behaviour, (Supernak et al., 2013; Yang and Sun, 2013; Afshari et al., 2021; Iryo-Asano et al., 2015) showing that as the crosswalk length increases, the signal violation behaviour significantly decreases. A one-meter increase in the crosswalk length reduces the likelihood of pedestrian violation by 13.9% (Afshari et al., 2021). Further, longer crosswalks force pedestrians to change speed more frequently as they must complete the crossing manoeuvre quickly in the limited green time in the presence of turning traffic. Drivers cannot predict this type of sudden change in speed, which can lead to safety hazards (Alhajyaseen and Iryo-Asano, 2017). Studies on pedestrian signals with countdown timers showed an adverse impact on crossing behaviour. Wanjing et al. (2015) reported that countdown signals adversely increase the possibility of collisions with unexpected right-turn vehicles and crossing vehicles or bicycles. Biswas et al. (2017) reported that signal heads with countdown timers were responsible for higher vehicular stop line violations and pedestrian signal violations that put pedestrians exposed to the risk of getting hit by a vehicle. Additionally, pedestrian crossing infrastructure such as physical layout, the presence of a refuge island, or pedestrian guardrail also significantly affects pedestrian conformity to the traffic signals (Xu et al., 2013).

(e) Built Environment: Environmental factors such as weather and lighting conditions can influence pedestrians' proclivity to run at red-light signals at a macro level. Li and Fernie (2010) reported that pedestrians' walking speed and red-light running rate were higher in cold and snowy conditions compared to warm weather and when the pavement surface was dry. Further, pedestrians were observed to be more careful when crossing in the dark and poor visibility conditions (Liu and Tung, 2014). Additionally, allowance of vehicle parking near the crosswalk was found to impact road crossing behaviour. Mukherjee and Mitra (2020) studied the impact of on-street parking near the crosswalks on risky road crossing behaviour and found a 30% increase in violations. Similar observations were also reported in France by Dommes et al. (2015) and Tom and Granié (2011). However, another study conducted in Athens, Greece (Yannis et al., 2013) contradicts the above observations, which revealed that pedestrians take more caution when vehicles are illicitly parked near the crosswalks.

2.2.2 Social Information

The literature highlighted that, similar to non-social information, humans respond to a wide range of social information disseminated from others (Cherry, 1978). In the transportation domain, social information can be defined as information or cues disseminated by road users, i.e., pedestrians or drivers (Faria et al., 2010). Studies have shown that an individual can obtain information from others for processing decisions. Based on the interpretation of the obtained information and its application scenario, such information can be advantageous or disadvantageous. Some of these studies are discussed below.

(a) Negative Implications of Social Information: Several studies highlighted that the social information gathered during the road crossing process could be adverse (Faria et al., 2010). An example of such a scenario is the sensation-seeking tendency among adolescents. Researchers have shown that sensation seeking among adolescents may be attributed to social information gained from peers. Rosenbloom (2009) reported that sensation seeking could be a significant predictor of reckless behaviour on the road. Zhang et al. (2016) studied sensation-seeking and pointed out that social risk-taking was the most prominent predictor for sensation-seeking, as adolescents have lower risk perception. School-going adolescents are likely to show 5-6 times more high-risk behaviour when their peers did not set a good example during road crossing (Holm et al., 2018). Faria et al. (2010) observed that neighbours' behaviour significantly influences the individuals' crossing decisions and further reported that a pedestrian's likelihood to start crossing a road increases 1.5 to 2 times if their neighbours had already begun crossing in the red-light phase. These findings indicate the crucial impact of social psychology on unsafe crossing decisions.

Socio-cultural variables (such as the degree to which a country is considered "developed" and demographic input) were found to influence pedestrian behaviour (Hamed, 2001). In the era of digital devices, researchers have pointed out another form of social feature that contribute to unsafe road crossing known as "digital distraction", increasing dangerous road crossing behaviour and reducing caution. Digital distraction can be in the form of mobile phone talking, texting, and headphone usage. Young adults (aged between 16 and 30 years) were subjected to more digital distraction-related injuries (Cooper et al., 2012; Nasar and Troyer, 2013; Vujanić et al., 2014). Pedestrians involved in digital distraction were less likely to perform cautionary behaviour, i.e., look left or/and right (Bungum et al., 2005; Pešić et al., 2016), less likely to wait for traffic to stop (Hatfield and Murphy, 2007), therefore increasing the crossing risk. Evidence from existing studies also suggested that digital distraction from electronic devices increased the likelihood of illegal crossing or signal violation behaviour of the pedestrian (Mukherjee and Mitra, 2020; Nasar and Troyer, 2013; Stavrinos et al., 2011; Zegeer and Bushell, 2012).

(b) Positive Implications of Social Information: In a social context, social cues obtained from another's actions (i.e., crossing a road) often play a pivotal role in one's crossing decision (Giraldeau et al., 2002). In human society, especially in groups, norms, beliefs, rules, and moral standards determine individuals' actions. Rosenbloom (2009), studied single versus group road crossing behaviour, stated that a group has the power of diffusion of responsibility and de-individuation as each group member feels that the responsibility of violating the norm is shared with the rest of the group members. Thus, individuals deter themselves from taking unsocial steps that violate norms, rules, and regulations in a social context. Similar social influence in road crossing was observed by Pele et al. (2017) in a comparative crossing behavioural study of Japanese and French pedestrians. The findings highlighted that the Japanese were more respectful of the rules than the French.

Further, people waiting for the green light were more committed to social order, hence, stuck with social norms (Wang et al., 2011). Tyler T.R. (1990) reported that the diverse types of motives (waiting/crossing) could be explained by the two perspectives of obedience to laws; one is called the instrumental perspective, and the other is called the normative perspective. People were influenced by gains and losses relating to obeying or disobeying the law as per the instrumental perspective. For example, time-saving might be one gain from violating the law or crossing in the red-light phase. The normative perspective explains the sense of duty to obey traffic law as a function of personal values.

2.2.3 Personal Characteristics

Demographic characteristics such as age and gender are found to impact the risky crossing decision. Most research found that male pedestrians were more likely to violate the red-light signal than females (Guo et al., 2011; Mukherjee and Mitra, 2020; Rosenbloom, 2009). According to Mukherjee and Mitra, the risk of pedestrian signal violations is significantly higher for people aged 16 to 49 in Kolkata, India (Mukherjee and Mitra, 2017). Brosseau et al. (2013) found that the age group of 18 to 35 has a strong tendency to violate signals in Canada. The pedestrians' trip purpose could also determine their road crossing behaviour. Zhang et al. (2016) reported that people who are on a stroll (walk for leisure) are more likely to have a safe crossing action than those who are heading to work or school, as they are friendlier with the conditions at that intersection. Further, studies showed that pedestrian with vision impairment faces crossing difficulties at intersection crosswalks. Bentzen et al. (2004) studied the crossing behaviour of blind pedestrians in Portland (USA) and revealed that blind pedestrians had considerable difficulty determining the onset of the walk signal phase, and more than 40% of them failed to begin crossing during the walk signal, resulted in late entry and ended up crossing the road in red (do-not walk) pedestrian-signal phase.

2.3 Pedestrian Distracted Road Crossing Behaviour at Intersections

To understand the impact of distraction on pedestrians' road-crossing behaviour, past studies have considered several types of distraction that may alter safe road-crossing behaviour. The common distractors were talking on the phone, texting, use of headphones for music listening or talking (Bungum et al., 2005; Praveen Kumar N. et al., 2018; Russo et al., 2018; Thompson et al., 2013; Wells et al., 2017; Zhang H. and Zhang C. and Chen F. and We IY., 2017), group talking (Praveen Kumar N. et al., 2018), and eating/drinking/smoking (Wells et al., 2017). A summary of recent distraction studies and the share of different distractions across different countries are reported in Table 2.1.

Past studies have pointed out that different types of distractions have different effects on safe road-crossing behaviour. There have been multiple studies in developed countries and few in developing nations on distracted pedestrian crossing behaviour. These past studies have been presented in two broad areas: (a) observational and (b) lab-based experimental studies.

2.3.1 Observational Studies

Naturalistic observational studies aim to inconspicuously monitor the pedestrians' road crossing behaviour in real-world settings (refer Table 2.1). The results suggested a range of unsafe behaviour; distracted pedestrians using their mobile phone was less likely to act in a cautionary way, i.e., look left and right (Bungum et al., 2005; Pešić et al., 2016); they are less likely to wait for traffic to stop (Hatfield and Murphy, 2007). Similar studies involving pedestrian distraction reported reduced

TABLE 2.1: Observed share of pedestrian distraction across different countries

Authors (year)	**Study Location**	**Location Type**	**Observations on % distraction**
Praveen Kumar N. et al. (2018)	India	Non-signalised intersection	Overall distraction (22.6%): mobile talking (11.3%), texting (0.6%), listening to music (3%) and group talking (7.7%).
Russo et al. (2018)	USA	Signalized intersections	Overall distraction (13.5%): mobile talking (2.9%), texting (5.7%) and headphones (3.7%).
Barin et al. (2018)		Intersection and Mid-block	Overall distraction (23%): mobile talking (3.4%), texting (8.5%) and headphones (12%)
Ortiz et al. (2017)		Intersections	Overall distraction (49%): mobile talking (27%), talking with others (46%), headphones (16%), eating and drinking (6%), and other (5%)
Wells et al. (2017)		Signalized intersection	Overall distraction (35%): mobile talking: (4.7%), texting (7.5%), headphones (19%) and eating and drinking (2.6%).
Violano et al. (2015)		Signalized intersection	Overall distraction (19%): Out of which mobile talking and texting (8%), head phones and ear buds (9%) and eating and drinking (2%).
Lin et al. (2015)		Intersections	*Overall distraction (19%)*
Thompson et al. (2013)		Signalized intersection	Overall distraction (30%): listening to music (11.2%), texting (7.3%), handheld phones (6.2%) and looking left/ right (34.9%).
Bungum et al. (2005)		Intersection	Overall distraction (20.8%): phone and headphone use (5.7%), Of the total observation: [looked left (54.8%), looked right (41.1%)].
Pešić et al. (2016)	Serbia	Non-signalised crosswalks	Phone use (10-15%)
Mohd Syazwan et al. (2017)	Malaysia	Signalized intersection	Overall distraction (25.8%): mobile phone use (84.8%), drinking/eating (4.5%), reading (4.8%) and smoking/talking (5.9%)
Chen and Pai (2018)	Taiwan	Uncontrolled intersection	Overall distraction (73.3%): mobile talking (10.2%), texting (15.2%), listening to music (6.6%), talking using an app (11.2%), web surfing (4.6%) and gaming (25.5%)
Zhou et al. (2019)	China	Signalized intersection	Overall distraction 7.7%: mobile talking (2.5%), texting (4.3%) and listening to music (0.9%)
Zhang H. and Zhang C. and Chen F. and We IY. (2017)		Non-signalised intersections	Overall distraction (15.1%): mobile talking (3.1%), watching mobile screen (11.1%) and listening to music (1.0%).

situational awareness and increased ignorance of surroundings and even caused inattentional blindness (Hyman et al., 2009). Thompson et al. (2013) analysed digital and social distraction and observed that texting pedestrians were 3.9 times more likely to exhibit at least one unsafe crossing behaviour. In contrast, pedestrians conversing in groups were 1.69 times more likely to exhibit unsafe crossing behaviour. A meta-analysis of the experimental effects of mobile phone conversations, texting, browsing, or listening to music on the behavioural measures of initiation duration, missed opportunities, crossing duration, glance (left and right), and hits and close calls showed that texting or browsing had the most detrimental effects causing hits and close calls (Simmons et al., 2020). Additionally, pedestrians who talk or text were subject to higher rates of near misses with vehicles. Talking on phones or texting might not significantly influence crossing speed, but it may increase startup time (mobile talking: 21%; texting 31%), (Gillette et al., 2016) and even crossing delay (Chen and Pai, 2018). It is also observed that as the waiting time at an intersection increases, most pedestrians are likely to use their mobile phones to kill undue halt time, and contrarily, a significant portion of them miss the adequate time of crossing (Zhou et al., 2019) and finish crossing late. Evidence from existing research also suggested that digital distraction from electronic devices increases the likelihood of pedestrian signal violation behaviour (Mukherjee and Mitra, 2020; Nasar and Troyer, 2013; Stavrinos et al., 2011; Zegeer and Bushell, 2012). Similarly, pedestrians who primarily engaged in texting during crossing were found to commit more crosswalk violations (Russo et al., 2018; Zhou et al., 2019).

In contrast to the unsafe behaviour displayed by mobile phone users overall, the finding of gender-specific differences was mixed. It was observed that female pedestrians were more likely to engage in mobile phone distraction compared to men (Baswail et al., 2019; Cooper et al., 2012). They were also less likely to look at traffic before crossing, and wait for traffic to stop (at signalised crossings); however, they were more likely to wait for a vehicle to stop at unsignalized intersections (Hatfield and Murphy, 2007). Another study on distraction reported no significant gender difference (Hyman et al., 2009). Further, a pedestrian's social status (single or in a group) during road crossing influences distractive engagement. Studies showed that pedestrians crossing by themselves were more likely to talk on their mobile phones compared to crossing in a group (Cooper et al., 2012; Russo et al., 2018). Notably, young adults aged between 16 and 30 years old were subject to more mobile phone distraction-related injuries (Cooper et al., 2012; Nasar and Troyer, 2013). Neider et al. (2010) explored the impact of distraction on older adults and found that older adults were more vulnerable to dual-task impairments, especially when conversing on their phones, than when listening to music or undistracted.

2.3.2 Experimental Observations

Experimental studies involve a controlled manipulation of different features, often to mimic the real-world environment, and investigate the impact of such changes on pedestrian road crossing behaviour. All such simulated studies have revealed several negative impacts of mobile-related distraction. Haga et al. (2015) investigated the impact of distraction in the virtual environment and reported a higher risk of accidents among pedestrians who use mobiles specifically for gaming. Horberry et al. (2019) also reported that pedestrians interacting with their phones were more likely to get hit by a vehicle, hit opposite-directional pedestrians, and even miss surrounding information. Based on an experimental investigation of outdoor crossing behaviour, Jiang et al. (2018) reported that mobile talking increased crossing time, while texting reduced visual attention span; similarly, listening to music increased the chances of crossing late. Lichenstein et al. (2012) reported sensory deprivation (especially auditory loss) due to distraction. The use of headphones with hand-held devices may present a safety risk amongst pedestrians in the presence of moving vehicles. Simulation-based experimental studies involving young participants have shown that the distracted pedestrians were relatively less successful in crossing a road (distracted: 80% vs undistracted: 84%), (Neider et al., 2010) and took more time (0.15 seconds more) to finish crossing (Banducci et al., 2016). Further, reading or texting while crossing a road contributed to the highest perceived task workload, rendering reduced roadside awareness (Banducci et al., 2016; Lin and Huang, 2017).

2.4 Pedestrian Waiting Behaviour at Signalised Intersections

The influence of waiting time in signal violation behaviour has been discussed in Section 2.2.1. Past literature showed that apart from demographic characteristics, variables such as timestamps of different events (arrival, departure, and completion), crossing characteristics (waiting duration, rolling effect, platoon and crossing speed), violation (signal and crosswalk), traffic characteristics (volume) and infrastructure (crosswalk length, signal phase length, width of refuse island) were extracted from video graphic survey and frequently used for analysing unsafe crossing behaviour. In the past decade, various studies have investigated the signal violation behaviour at intersection crosswalks using various statistical approaches. A significant number of them relate to the waiting time with the associated risks during the crossing. Various modelling approaches have been used to model the signal violation behaviour. The most commonly adopted method has been the logistic regression (refer Table 2.2) to estimate the violation probabilities and classify whether a pedestrian will wait or cross in red-light phase (Rosenbloom, 2009; Yang and Sun, 2013; Yang et al., 2015).

TABLE 2.2: Various empirical study details focused on pedestrian signal violation behaviour

Authors and Country	Type of Study	Type and number of sites studied	Sample Size	Variable used	Model/ method used
Afshari et al. (2021), Iran	V	Crosswalk (10)	1590	Traffic and pedestrian volume, the number of violators, length of the crosswalk, red light duration, physical movement problems, platoon, waiting time and existence of a safe place after violation.	Binary logistic regression
Dhoke et al. (2021), India	V	Signalized intersection crosswalk (4)	2653	Gender, age, luggage, pace, group size, arrival and departure signal, crosswalk use and crossing speed	Kaplan-Meier survival curve, Cox Proportional Hazard model and Accelerated Failure Time model
Mukherjee and Mitra (2020), India	V & Q	Signalized intersection crosswalk (55)	V: 65500; Q: 3250	Gender, age, traffic and pedestrian volume, crosswalk and signal violation, walking speed, speed change, path change, rolling, waiting time, overhead load and distraction (yes/no)	Beta regression
Russo et al. (2018), US	V	Signalized intersection crosswalk (4)	3038	Gender, age, group size, crossing from opposite direction, arrival and departure time, waiting time, pedestrian signal on start and end of crossing, push button, distraction (talking, texting, headphones and others)	Binary logistic regression
Cao et al. (2017), China	V & Q	Crosswalk (4)	1086	Gender, age, direction of crossing, width of the refuge island, signal position, number of pedestrian waiting, group size, entrance and exit flow, V/C ratio and conflict	Random-effects ordered logistic regression and binary logistic regression
Chen et al. (2017), China	V	Crosswalk (13)	1075	Gender, age, arrival and departure time, waiting time, crosswalk length, refuge island, conflicting traffic volume, clearance time, red time and countdown	Binary logistic regression
Biswas et al. (2017), India	V	Signalized intersection crosswalk (2)	943 with & 860 without timer	Gender, age, arrival and departure signal, countdown timer, cycle time, crosswalk length, violation, waiting time and arrival pattern	Non-linear regression
Alexandre and Ruggiero (2015), UK	V	Signalized crosswalk (1)	2455	Gender, luggage, group size, waiting time, location of the sidewalk, crossing direction, crossing in two steps, traffic, gap size, vehicle type, vehicle lane, number of pedestrians violated the signal since arrival and number of pedestrians waiting at sidewalk	Binary logistic regression
Yang et al. (2015), China	V & Q	Signalized intersection crosswalk (5)	V: 1181	Gender, age, arrival time, departure time, left-turn phase upon arrival, traffic light status, traffic volume, number of people waiting upon arrival and number of people crossing upon arrival	Logistic regression and Accelerated Failure Time model
Brosseau et al. (2013), Canada	O & V	Crosswalk (13)	2518	Gender, age, number of pedestrians waiting, arrival and departure time, group size, crossing speed, violation and dangerous crossing, and waiting time	Logistic regression

Table 2.2: Continued from previous table

Authors and Country	Type of Study	Type and number sites studied	Sample Size	Variable used	Model/ method used
Yang and Sun (2013), China	V & Q	Signalized crosswalk (10)	V: 1518, Q: 356	Signal violation, waiting time, red and green time, crossing distance, crossing lanes, time gap, space gap, vehicle speed and vehicle volume	Logistic regression
Guo et al. (2011), China	V & Q	Signalized crosswalk (7)	V: 1497, Q: 356	Gender, age, safety awareness, group size, conformity psychology, red-time, safety gap, pedestrian and traffic volume, travel time and trip purpose	Kaplan-Meier estimate, Cox Proportional Hazard regression
Tom and Granié (2011), France	V	Crosswalk (4)	400	Tempo while approaching curb, number of pedestrians waiting, tempo at curb, before and during crossing glance, departure signal, starting and ending on zebra markings, and type of crossing	Kaplan-Meier estimate and Cox Proportional Hazard regression
Faria et al. (2010), UK	V	Crosswalk (1)	365	Social information and neighbour behaviour	Nearest neighbours
Rosenbloom (2009), Israel	O	Signalized crosswalk (1)	1392	Gender, age, number of pedestrian waiting and joining on red, someone crossing in the red-light phase and traffic	Binary logistic regression
Tiwari et al. (2007), India	V	Signalized intersection crosswalk (7)	827	Gender, age, waiting time, signal violation, cycle time, green and red-phase, delay and crosswalk width	Survival analysis (Kaplan-Meier estimate)
Yagil (2000), Israel	Q	Signalized intersection crosswalk (7)	205	Gender, age, normative motive, situation motive, instrumental motive and situational factors	Regression

Note: V: Video graphic survey, Q: Questionnaire survey, O: Observational survey

Even though this method is useful for identifying the significant factors, it does not account for the heterogeneity of the waiting time factor, which is observed to be an essential factor in signal violation behaviour. Recently, safety researchers started applying hazard-based duration models to explore pedestrians' street-crossing behaviour. A hazard-based duration model is a probabilistic approach that is well suited to analyse time-to-event observations. Pedestrian crossing behaviour has been studied using non-parametric, semi-parametric, and parametric models. For example, Tiwari et al. (2007) explored the pedestrian's probability of risk exposure at signalised intersections in Delhi (India) using the non-parametric Kaplan–Meier (KM) process. However, it did not reflect the effects of various internal and external factors. Hamed (2001) also studied pedestrians' violation risks and waiting times at intersections in Jordan using fully parametric hazard models. The study, however, did not consider the censored data on waiting times. Guo et al. (2011) investigated pedestrian red-light crossing activity in China using semi-parametric and fully parametric hazard models to determine the effects of personal characteristics, traffic characteristics, and trip features on signal violation behaviour. Similarly, Dhoke et al. (2021) scrutinised the impact of waiting time on violation at three signalised intersections and reported that the Weibull Accelerated Failure Time (AFT) model could be best suited for modelling violation behaviour.

2.5 Pedestrian-Vehicle Interaction at Signalised Intersections

As per the literature discussed in the section 2.2, signal violation compromises pedestrian road crossing safety. Studies showed that crossing safety is further compromised when pedestrian starts accepting unsafe gaps against fast-moving traffic. Safety margin estimation remains the most popular dummy measure to quantify the risky crossing behaviour at different crossing locations across road networks. A variety of slightly deferring definitions have been adopted by different researchers (Chu and Baltes, 2001; Lobjois and Cavallo, 2007; Jain et al., 2014). However, fundamentally, they wish to convey a common interpretation that if the pedestrian is slower by a time equal to the safety margin, the conflict is bound to happen. The time interval between the instance at which the pedestrian crosses the intersection point of the path of the pedestrian and the path of a vehicle, also called the conflict point, and the instant at which the next vehicle crosses the point is called the safety margin. This measures the level of safety or the extent of risk a pedestrian is willing to take under existing conditions.

To work out the effect of location on the safety perception of pedestrians, safety margin studies have been conducted on different sections of roads. Onelcin and Alver (2017) and Koh and Wong (2014) have considered signalised intersections as study areas in their respective cities. Many studies have also been conducted at unsignalized intersections (Boroujerdian and Nemati, 2016; Vasudevan et al., 2020). Mid-block locations, being prone to comparatively severe accidents, have been given special attention, and many studies are based on this (Kadali et al., 2015; Chaudhari et al., 2020).

2.5.1 Safety Margin Models used in Past Studies

For safety margin and gap acceptance studies, the most common model used till now is Multiple Linear Regression (MLR), where many variables can be included in a single model simultaneously, and their coefficients help to translate the effects on the model (Shaaban et al., 2019; Kadali et al., 2015; Kadali and Vedagiri, 2013). With the advancement in soft computing techniques, an advanced model such as Artificial Neural Network (ANN) model was developed by researchers to overcome the limitations of linear regression (Kadali et al., 2015). The non-linear nature of the ANN model is efficient in predicting the effects of variables on gap acceptance in mixed traffic conditions but lacks model interpretability. However, MLR is easier to quantify the effects of significant variables on gap selection. Further, the log-normal regression model has also been used by researchers to examine how various variables play their part in pedestrian gap acceptance (Yannis et al., 2013).

2.5.2 Factors Influencing Safety Margin

Several factors have been known to impact the safety margin left by pedestrians and the gaps they accept, contemplating them to be safe for crossing the streets. These factors can be broadly grouped into pedestrian demographic characteristics, behavioural characteristics, vehicular characteristics, traffic characteristics, and environmental factors.

(a) Pedestrian Demographics: Pedestrian demographics significantly shape the way people react to signals, which strongly correlates with the safety adopted by them. Numerous researchers have come up with studies demonstrating how these factors drive pedestrian behaviour. Gender is one of these factors. Males are mostly seen as risk-takers than females, (Mukherjee and Mitra, 2020; Rosenbloom, 2009; Tiwari et al., 2007; Moyano Diaz, 2002). Pele et al. (2017) quantified this result through their study in France as well as Japan, indicating that men are more likely to cross a red light (40.6% breaking of rules) than women (correspondingly 25.7%). Zhu et al. (2021) conducted a study in Hong Kong and found that males are more likely to run red-light compared to females and expose themselves to smaller safety margins (Zhuang and Wu, 2011, 2012). Onelcin and Alver (2017) observed that the mean safety margin of males (7.43 s) had been found to be lower than females (8.25 s). But they did not find gender to be a significant variable in determining the safety margin. These findings are also in line with the inferences drawn by Boroujerdian and Nemati (2016).

Age has been found to be the major factor impacting pedestrian safety (Chu and Baltes, 2001; Oxley et al., 2005; Lobjois and Cavallo, 2007; Holland and Hill, 2007). Experiments conducted by Oxley et al. (2005) in a virtual environment reported that young people are capable of making judgements based on distance, as well as time gap and vehicular speed, which are a little difficult for older people to process in a limited time. Contrasting results were obtained in a similar study by Lobjois and Cavallo (2007), where all the age groups seemed to take an equal risk (5.8% for 20 to 30, 8.2% for 60 to 70, and 5.9% for 70 to 80 age groups). Here, it appeared as if older people were aware of their slow mobility. They selected a larger time gap compared to other groups.

Group size is another important demographic factor. It denotes the number of pedestrians together. Henceforth, whether waiting and crossing alone or in groups also has an impact on illegal crossing. In a study by Rosenbloom (2009) in Israel, she found that pedestrians waiting alone are more likely to end the wait and cross- compared to those waiting with other people. This can be attributed to the theory of social control. Similar observations were made by Koh and Wong (2014) & Brosseau et al. (2013). It was concluded more conclusively by Zhuang and Wu (2012) through a study in Hangzhou, China, that the safety margin was higher for pedestrians in bigger groups. As the group size grows, pedestrian flow increases and drivers become more alert, approaching these groups at lower speeds, thus improving safety. Even in India, Avinash et al. (2020) supported this conclusion, and as per them, an increase in group size adds to safety, due to the alertness of drivers and subsequent yielding.

(b) Pedestrian Behavioural Characteristics: Similar to pedestrian demographics' role in their safe crossing characteristics, their behavioural characteristics are equally pivotal. Behavioural attributes such as pedestrian speed, waiting time endurance, looking towards vehicles before and during crossing, distractions, and the nature of crossing (perpendicular or oblique) are the chief influencers. The crossing speed of male pedestrians (1.168 m/s) is almost always higher than female pedestrians (1.061 m/s) (Chaudhari et al., 2020). This has been in line with the study of Onelcin and Alver (2017), who emphasised crossing speed exclusively, and found that females, people with luggage, older pedestrians, and people in groups of two or more had lower crossing speeds. While studying 254 pedestrians in Hangzhou, China, Zhuang and Wu (2011) concluded that the safety margin tends to follow a log-normal distribution, indicating that majority of people leave a smaller safety margin while crossing. Running people mostly fall in this category. This means that higher crossing speeds lead to reduced safety for pedestrians. Inadequate reaction time available for drivers to yield is the main reason behind it. Chaudhari et al. (2020) stated that pedestrians maintain higher speeds when crossing in oblique directions than when crossing perpendicularly. More specifically, people tend to leave smaller margins when crossing in an oblique manner.

Another important aspect of pedestrian behaviour is their frequency of looking at vehicular traffic before and during the crossing. When a person is looking at the traffic more often, it suggests that he/she is more concerned about safety. This has been depicted by Zhuang and Wu (2012) through their study in China that a higher frequency of looking before crossing meant that pedestrians are exhibiting safer behaviour which increased safety. This happens partly by reducing the running frequency and returning back to look for safer gaps.

Endurance to the waiting time explains the critical gap from the approaching vehicles accepted by pedestrians in order to avoid conflicts. In a study in Athens, Yannis et al. (2013) found that if people get a time gap of 6 seconds or more, the crossing probability is almost 100%. In Singapore, Koh and Wong (2014) found that people were more cautious when violating the red signal at the near end than at the median, and usually accepted longer time gaps (6.3 s) at the near end than when at the median or far end crossing (5.2 s). This means that once they start crossing by accepting a gap, there is an urge to finish the task and feel safe at the other end; and in the process, they are open to accepting smaller gaps in the second phase. A similar observation was made by Zhao et al. (2019), who also explained the effect of waiting time on gap acceptance. More than 50% of the pedestrians are expected to cross after a waiting time of 33 seconds while at the kerb side, whereas the same proportion of people could wait only for 12 seconds before they start accepting gaps when they have reached the median. At the same time, the probability of accepting the available gaps increases with the increase in waiting time whether the pedestrian is at the kerb or at the median, because of the diminishing patience of those waiting for their turn to cross. Hence, gap size decreases with an increase in waiting time (Kadali et al., 2015).

The usage of handheld media has escalated in recent times and has driven a large proportion of the public to hinge on these continuously. This has had its consequences. Pedestrians are no exception. Their behaviour is significantly modified. Hence, this has been one of the factors researchers have keenly studied. Distractions, in the form of using mobile phones, and also because of smoking, drinking and eating, individually or coupled, have proven to be setbacks for the safety of pedestrians, according to Zhuang and Wu (2012). An interesting observation by them was that pedestrians who settled their hair or clothes during crossing had a bigger safety margin than those who did not. Vasudevan et al. (2020) analysed that when distracted by smartphones or other devices, the critical gap of pedestrians increased. These distractions can leave pedestrians busy or engaged, and prevent them from getting impatient and choosing dangerous gaps. This was not consistent with an earlier study by Schwebel et al. (2012), where, in a virtual street environment, they found that distracted pedestrians had almost equal safety margins as that of undistracted ones. Nevertheless, college students texting or listening to music while crossing the street were hit more often by vehicles, thereby having their safety compromised.

(c) Vehicular Traffic Characteristics: The safety of pedestrians is not solely dependent on them. The running traffic also has a big role in it. The effect multiplies when there is poor lane discipline in traffic, which is mostly the case in developing countries. Vehicle type is one of the most crucial determinants for safety margin (Zhuang and Wu, 2011, 2012). Interactions with heavier vehicles lead to pedestrians accepting larger gaps. People maintain less margin for two-wheelers and three-wheelers as compared to heavy vehicles. In Malaysia, Hamidun et al. (2016) marked that in all situations, higher approaching speed presents a higher potential of an accident; hence it is a nemesis of safety for pedestrians. In a simulation-based study in France, Lobjois and Cavallo (2007) found that the safety margin of elderly people reduced with an increase in vehicular speed. However, in general, as per Avinash et al. (2020), as speed increases, the safety margin also increases, which may be due to the pedestrians increasing speed accordingly. But this may lead to more severe accidents. Similar conclusions were made by Demiroz et al. (2015). Further, Pawar and Patil (2015) observed that the speed of the conflicting vehicles was significant in the spatial gap but not in temporal gap acceptance.

(d) Environmental Condition: Environmental factors include both the natural (time of day and weather conditions) and man-made (road geometry, crosswalk markings, signals, land-use, etc.) settings around which pedestrian-vehicle interaction is studied. Past studies revealed that day of the week, time of the day influences pedestrian crossing behaviour. Liu and Tung (2014) observed that the safety margin at dusk was remarkably larger than that at midday. This shows cautionary behaviour from the pedestrian's side when the visibility is low. Road geometry is another influencing factor which may increase or decrease the safety of a crossing pedestrian. Chaudhari et al. (2020), in their study of urban mid-block road sections, found that pedestrians were increasing their speed in the farther lane to reach the median (safe zone) quickly. Similar behaviour is also observed during the second stage (from the median to the other end of the road) of crossing, which shows the urge to finish the crossing process to reach the other end, the so-called safe zone. Further, vehicles parked illegally near the vicinity of the crosswalk were found to influence pedestrian crossing behaviour. Yannis et al. (2013) observed that illegally parked vehicles drove pedestrians to choose larger gaps, discouraging them from taking risks while crossing. Here, hindered visibility seemed to make them more cautious. In contrast to the observation made by Yannis et al. (2013), another study conducted in Kolkata by Mukherjee and Mitra (2019) highlighted an increase in risk-taking behaviour in the presence of parked vehicles.

2.6 Research Gaps from the Literature Review

The existing studies on pedestrian signal violation, distracted road crossing, waiting behaviour and safety margin were discussed in this chapter. The identified research gaps from the literature review are listed as follows:

- Only a handful of studies were conducted in Lower and Middle-Income Countries (LMICs) to understand the pedestrian risk-taking behaviour at signalised intersection crosswalks, India in particular.
- A limited number of papers addressed the impact of non-social factors (such as signal phase length and waiting time) in India. Studies investigating the impact of social factors on risky crossing behaviour were conducted in developed nations, considering a smaller sample size, which necessitates further exploration in LMICs, especially in India.
- Several studies have been carried out in developed nations displaying the potential impact of distracted road crossing behaviour on crossing safety risk.

However, their conclusions were not validated with accident statistics due to the paucity of such data availability/inventory. Ralph and Girardeau (2020) raised similar concerns and pointed out that despite the lack of field evidence in terms of accidents data, scholars tend to describe distracted walking as a severe problem, which might misplace the key focus from the overall effort of improving safety for all road users. Therefore, in the absence of distraction-related accident records in LMICs, a questionnaire survey of a representative sample could provide an overall accurate initial estimate of pedestrians with distraction and its severity in daily commutes (while in a road environment).

- Majority of the researchers investigated the impact of digital distraction on unsafe crossing behaviour, and very limited studies considered the influence of social distraction (group talking).

- The traffic glance and its importance in road crossing behaviour has been extensively studied in the context of distracted road crossing, but very few studies investigated its impact on pedestrians' waiting behaviour at signalised intersections. The unsafe behaviour displayed by pedestrians while crossing the road under distraction (digital and social distraction) was not extensively analysed to understand the pedestrian waiting behaviour, especially using the time-to-event analysis (survival analysis) approach. This raises the future scope for investigation.

- Researchers have explored pedestrian gap acceptance (in terms of safety margin) extensively on mid-block sections. The majority of the studies were conducted considering common pedestrian and traffic level attributes. A limited number of studies investigated the impact of pedestrian level attributes such as crossing speed, traffic glance and distraction, and traffic level attributes such as type of vehicle, location of gap acceptance, the direction of approaching vehicle and number of traffic plying on the road.

2.6.1 Summary

In this chapter, most of the existing studies related to pedestrians' road crossing behaviour were discussed and presented. The literature review was carried out in four stages.

In the first stage, existing studies on pedestrian signal violation behaviour were reviewed and also highlighted the influence of social and non-social factors at signalised intersection crosswalks.

In the second stage, the studies related to pedestrian distraction were reviewed. The influence of different types of distractors and their influence on safe/unsafe road crossing behaviour was presented. It was also been highlighted those past study findings were not validated with accident statistics due to the paucity of such data availability/inventory in developing countries.

In the third stage, pedestrian waiting behaviour at signalised intersection crosswalks was presented. Existing models studying pedestrian waiting behaviour were reviewed in this chapter. Further, studies that used a time-to-event analysis approach were evaluated, and their gaps were presented.

In the fourth stage, studies that quantified the risky crossing behaviour using the safety margin approach were reviewed. From the literature review, a variety of critical setbacks were identified and quantified.

For solving the stated problems, a suitable study framework is proposed in the following chapter. To address some of the research gaps highlighted in Section 2.6, a detailed systematic study was conducted on the evaluation of pedestrian signal violation behaviour at signalised intersection crosswalks.

Chapter 3

Study Methodology

3.1 Introduction

Studies related to pedestrian risky crossing behaviour at signalised intersections are very limited in Lower and Middle-Income countries, especially in India. Inappropriate standards and poor maintenance of the foot over bridges and underpasses forced pedestrians to cross at grade level. As most pedestrians prefer or are obliged to use crosswalks instead of foot over-bridges, the inadequacy of crosswalks in terms of design or dimensions has reduced pedestrian safety. In addition, unoptimised and extended red-phase lengths and higher motor vehicle priority at intersections add to the inconvenience for pedestrians that force them to take unnecessary risks. Existing deficiencies in the infrastructure point out that more understanding of pedestrian risky crossing behaviour at intersection crosswalks is required to propose effective interventions and improve road crossing safety. This chapter presents the details of the study methodology used for developing the pedestrian signal violation model, distraction model, waiting behaviour model and safety margin model.

3.2 Overall Study Framework

To improve pedestrian road crossing safety, understanding their road crossing behaviour at signalised intersections is essential. To achieve the objectives discussed in Chapter 2, a simple research methodology is developed and presented in Figure 3.1

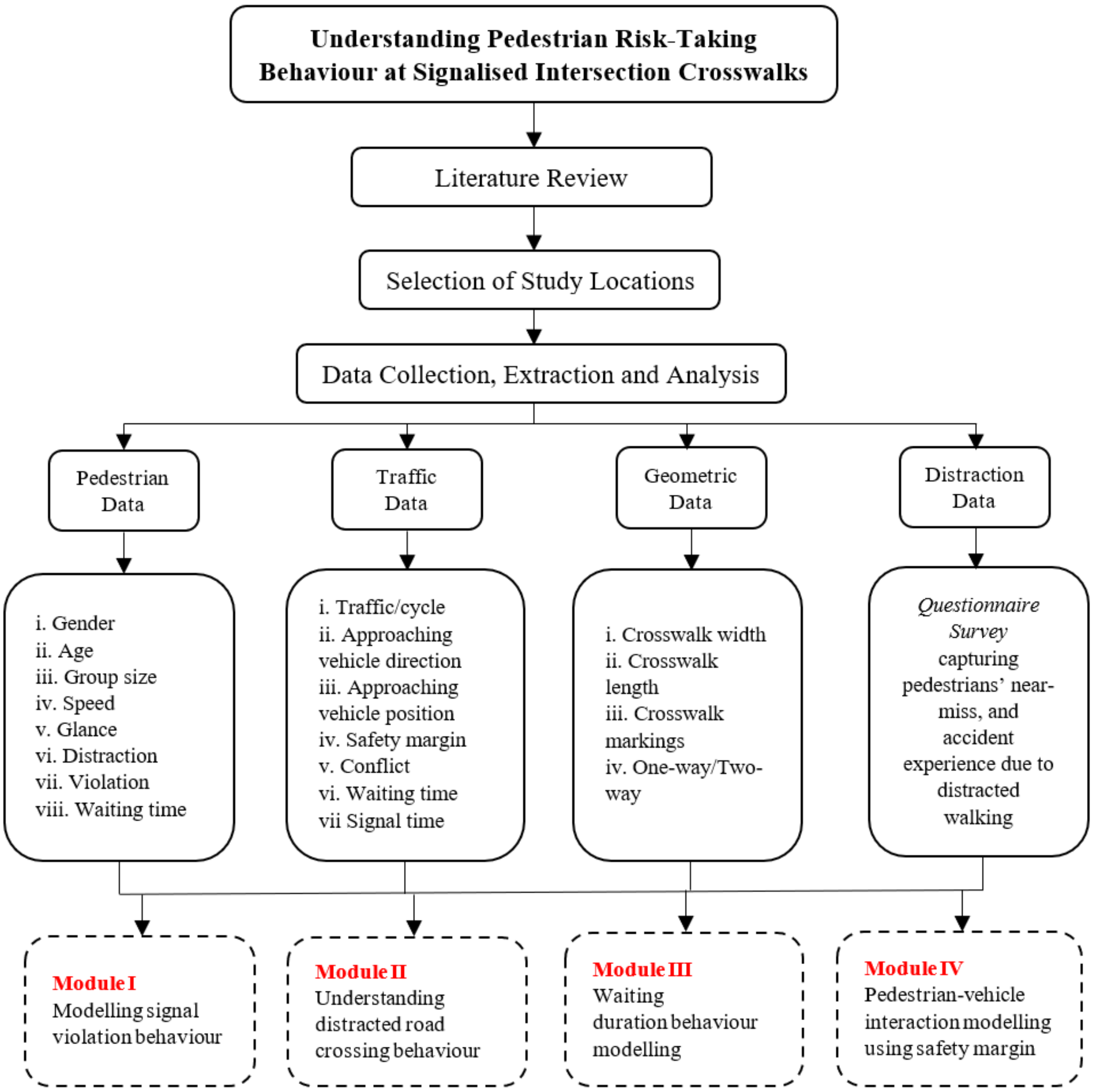


FIGURE 3.1: Flow chart of study methodology.

The research begins with a review of the existing literature to identify the state-of-art in modelling pedestrian signal violation, distracted road crossing, waiting behaviour and safety margin. Consequently, real-world field data was gathered from three one-way and eight two-way intersection crosswalks located in Kolkata, India. The collected data were extracted in a laboratory setting by trained research associates. Next, the extracted data were used to develop models. The model development and analysis are divided into four modules. In the first module, the pedestrian signal violation behaviour model was developed to understand the risk-taking behaviour at intersection crosswalks, the second module presented pedestrian distracted road crossing behaviour, and the third module explained the pedestrian waiting behaviour using survival analysis. The fourth module quantified the pedestrian crossing risk exposure using the safety margin approach.

Overall study framework illustrated in Figure 3.2 to 3.5 are briefly described in this chapter. Detailed explanations are provided in the individual chapters.

3.2.1 Selection of Survey Location

For site selection, initially, a preliminary survey was conducted in Kolkata city across the north and south stretch. The intersections with high pedestrian and traffic flow, marked crosswalks, and pedestrian signal heads on both sides of the road were given priority. A detailed description is provided in Chapter 4. Three one-way and eight two-way crosswalks are identified for the video data collection and to conduct a distraction-themed questionnaire survey. To capture the variability across Kolkata, sites were selected randomly across the north and south stretch. All the intersection crosswalks were installed with time-dependent signals where the signal's phase length was updated by the traffic regulator based on traffic flow. All the selected intersections are situated in the urban region of Kolkata city. The one-way crosswalks were located at Central Business District (CBD) location, and two-way crosswalks were located in mixed zones like residential, educational and/or commercial. The two-way crosswalks had no dedicated refuse island, but pedestrians can stand there in an unsafe manner while waiting to cross safely (i.e., from median to end).

3.2.2 Data Collection, Extraction, and Analysis

Data Collection

The data collection process requires careful planning to gather the required data for study analysis and model development. The data collection involved capturing pedestrian and traffic movement at intersection crosswalks precisely. Video graphic data were collected at the selected intersections to reduce the tedious manual data collection process. Simultaneously, a distracted-themed questionnaire survey was conducted to gather evidence on pedestrians subject to near-miss and accidents due to mobile use in road environments. In addition, geometric data were collected using the measuring wheel.

Data extraction

Data were extracted in a laboratory setting by well-trained research associates using video editing software. A total of 2360 pedestrians and their crossing behaviour were observed in one-way crosswalks. Additionally, 2800 pedestrian crossing behaviour was observed from two-way crosswalks. Further, 570 questionnaire survey responses were collected from eight survey locations. All possible pedestrian and traffic characteristics were extracted from the data to achieve the stated study objectives.

Data analysis

The extracted data were used for modelling pedestrian signal violation, distraction, waiting time and safety margin at signalised intersection crosswalks. The primary exploratory analysis and summary statistics were estimated to get an important insight into the data characteristics, which helped to choose appropriate analysis methods for further analysis. The following tools were used for the entire data analysis and model development process:

- **R Statistical Programming Language (version 4.1.1)**: R programming language was utilised for Exploratory Data Analysis (EDA), summary statistics estimation and model development (multiple linear regression and logistic regression). The following libraries have been utilised for the data analysis:
  - ***dplyr***: Data manipulation (data preparation).
  - ***ggplot2***: Data visualisation.
  - ***forcat***: Factor data manipulation.
  - ***margins***: To compute Average Marginal Effects (AME).
  - ***nagelkerke***: To estimate McFadden pseudo R-squared measure, along with p-values, for estimated models.
- **Python Programming Language (version 3.8)**: Python language was used to model the time-to-event analysis. The following libraries have been utilised for the data analysis:

- ***numpy***: Array data manipulation.
- ***pandas***: Data manipulation (data preparation).
- ***matplotlib and seaborn***: Data visualisation.
- ***lifelines***: Survival analysis (COX Proportional Hazard and Accelerated Failure Time model estimation).

- **Stata 17**: Stata was used for model variable selection, generating marginal plots, Kaplan-Maier (K-M) curve and survival plots for best fitted Accelerated Failure Time (AFT) model.

3.2.3 Methodology for Modelling Pedestrian Signal Violation Behaviour

Enhancing pedestrian convenience and road crossing safety requires an understanding of pedestrian unsafe road crossing behaviour. Understanding the question of why a pedestrian takes risk or what force a pedestrian to act in a certain way would help policy-makers and planners to propose interventions that would make road crossing safe and comfortable. The current understanding of pedestrian risky road crossing behaviour especially, the red-light crossing behaviour was extensively studied in developed countries considering non-social attributes of the transportation system such as infrastructure and traffic.

Only a few studies focused on the aspect of social factors that encourages unsafe crossing behaviour. Hence, this section has been undertaken with the aim of developing a suitable model accounting for various social and non-social factors to understand pedestrian signal-violation behaviour at signalised intersection crosswalks. Data extraction, including both social and non-social factors, is a time-consuming task. Thus, the current study is conducted based on the data gathered from three one-way crosswalks only. The one-way crosswalks were selected to simplify the data extraction, analysis, and understanding of the outputs. The complete methodology of the signal violation model has been illustrated in Figure 3.2.

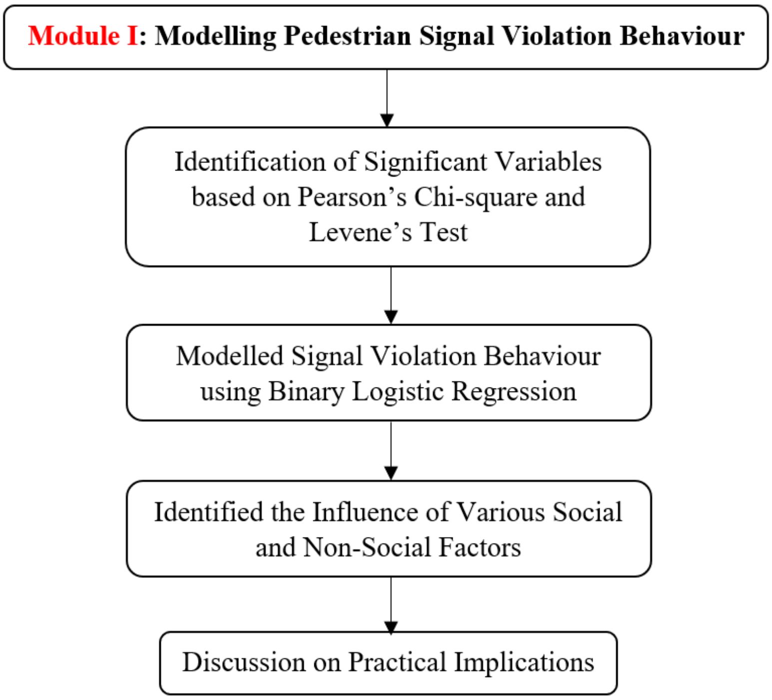


FIGURE 3.2: Module I – Analysis framework for modelling pedestrian signal violation behaviour.

In the current study, two types of data were extracted. One is related to non-social factors such as signal phase and cycle length to compute the waiting duration for safe crossing. The other is related to social factors, such as how many pedestrians were waiting on the subject pedestrian's arrival and how many pedestrians joined later. Further, the number of pedestrians crossing during the red-signal phase from the same and opposite directions was also observed. The estimated model presented in this study divulges the importance of social and non-social factors in pedestrian safe/unsafe road crossing decision-making.

3.2.4 Methodology for Modelling Pedestrians' Distracted Road Crossing Behaviour

The past studies cited that pedestrian risk-taking behaviour at intersection crosswalks is one of the major factors that contribute to a large number of fatal pedestrian-vehicle crashes. The improper pedestrian infrastructure has been cited as one of the reasons that forced pedestrians to take such risks. In the 21^{st} century, due to the digital revolution, mobile phone use has increased, especially among the young population. This led to another type of risk-taking behaviour that has reduced road crossing safety at intersection crosswalks. Mobile phone dependency has proliferated the use of such devices in the road environment, even when crossing through intersection crosswalks. This led to many injuries and admission to the ERs (Emergency Departments). Statistics revealed that developing country like India possesses more than 50% young (<25 years old) population (Commissioner, 2011; Ministry of Statistics and Programme Implementation, 2017). Thus, it is expected that parallel to traits observed in developed nations, young Indian pedestrians, also prone to excessive mobile use, perform similar unsafe crossing behaviour. Thus, proper investigation of distracted road crossing behaviour has to be done to quantify the behavioural differences between pedestrians with distraction and without distraction.

This study investigates the influence of different types of distraction on road crossing behaviour at signalised intersection crosswalks. The complete methodology of distraction models has been illustrated in Figure 3.3. In the absence of accident records-based validation, which is often noticeable in developing countries due to the lack of data inventories, a two-step approach has been adopted here. Initially, a field-based distraction-themed questionnaire survey was conducted at eight locations across Kolkata city (India) to investigate the severity of pedestrians' distracted road crossing in terms of near-misses and traffic collisions. Based on the evidence, a video graphic field-based observational study was proposed, and data was gathered from three signalised intersection crosswalks (one-way). In addition to digital distractions such as mobile phone talking, texting, using headphones, and holding a phone in hand, which have been extensively studied in the past (especially in developed countries), the present study also investigated the influence of social distraction (like group conversation) to understand the type of distraction (digital vs social) which encourages unsafe road crossing behaviour using binary logistic regression models.

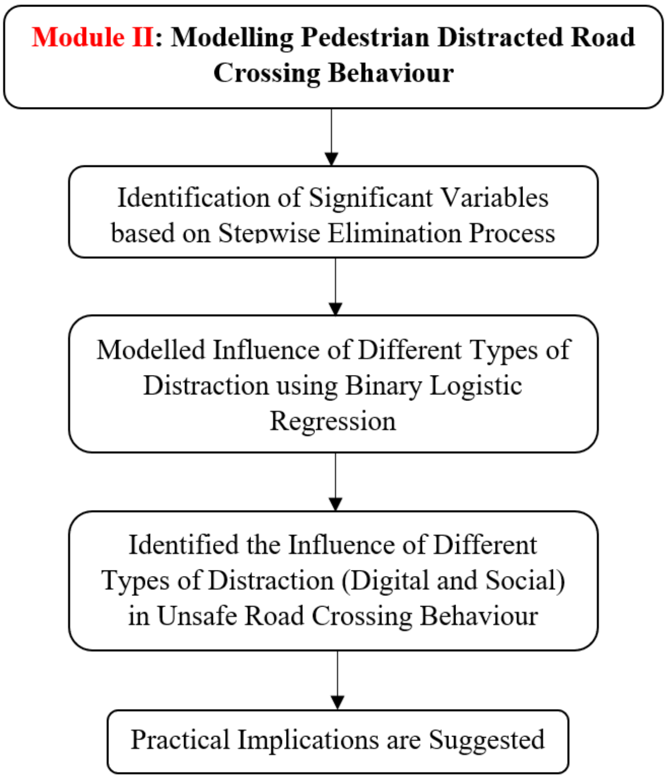


FIGURE 3.3: Module II – Analysis framework for distraction modelling.

3.2.5 Methodology for Modelling Pedestrians' Waiting Behaviour

In the past decade, various studies have investigated the signal violation behaviour at intersection crosswalks using various statistical approaches. A significant number of studies relate the waiting time with the associated risks during the crossing. Various modelling approaches have been used to model the signal violation behaviour. The most commonly adopted method was logistic regression to estimate the violation probabilities and classify whether a pedestrian will wait or cross in the red-light phase. Even though this method is helpful in identifying the significant factors, it does not account for the heterogeneity of the waiting time factor, which is an essential factor in signal violation behaviour. Thus, the current study investigates the waiting duration of pedestrians using time-to-event analysis (survival analysis) using data collected from eight two-way signalised intersection crosswalks from Kolkata. The complete methodology of the waiting duration model using survival analysis has been illustrated in Figure 3.4.

In this study, initially, a Kaplan-Meier curve (non-parametric) was plotted using the estimated survival probabilities (waiting), which gives an overview of the duration to which a certain proportion of observed pedestrians will wait for the oncoming green light. Though the KM estimate is informative, it still can't be used for checking the impact of covariates. To cope with the problem, a semi-parametric Cox Proportional Hazard model was used to understand the different factors influencing waiting behaviour. However, the proportional hazard assumption was not satisfied. Therefore, the parametric Accelerated Failure Time (AFT) model was used to understand the covariates affecting the waiting duration. Based on the results, several practical interventions were proposed.

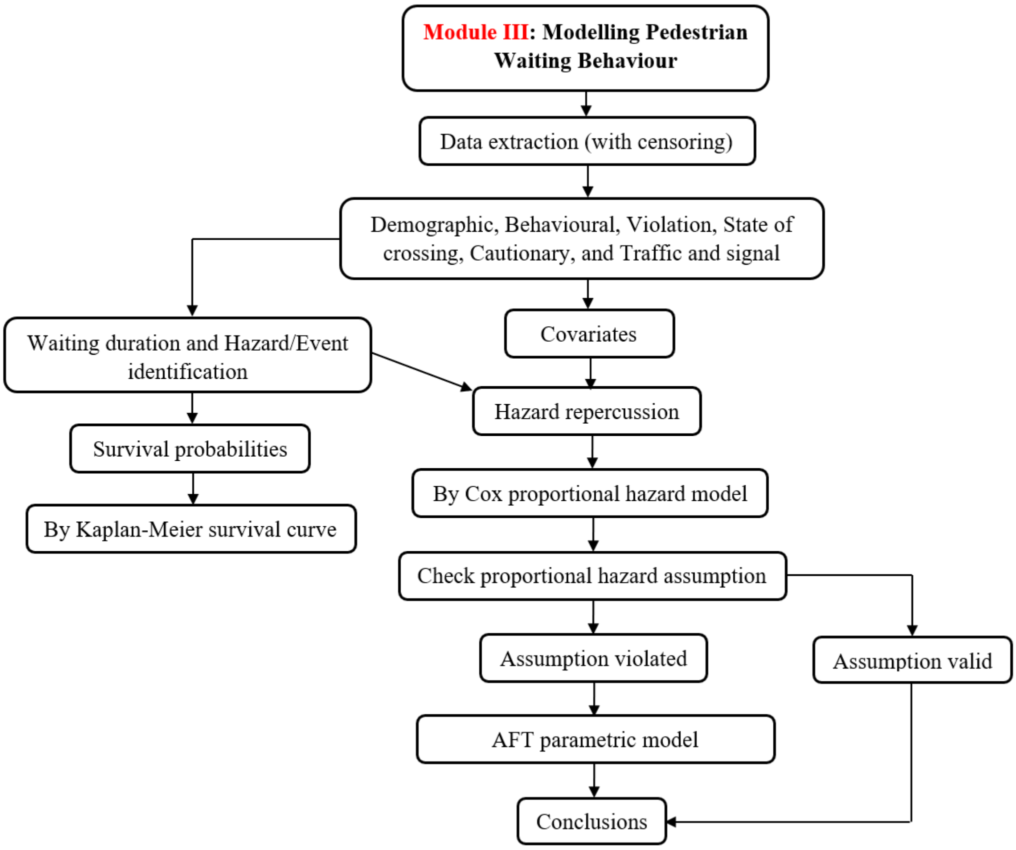


FIGURE 3.4: Module III – Analysis Framework for waiting duration modelling.

3.2.6 Methodology for Modelling Pedestrian Safety Margin (PSM)

Pedestrians exhibit higher risk-taking behaviour while crossing the road compared to walking on the sidewalk. The pedestrian-vehicle conflict measure is a useful technique to estimate different types of pedestrian-vehicle conflicts with variations in pedestrian and driver behaviour at signalised intersections. In this view, the safety margin is an appropriate technique to evaluate pedestrian road crossing safety at signalised intersection crosswalks. The Pedestrian Safety Margin (PSM) is defined as the marginal value maintained by pedestrians while accepting the vehicular gap during the road crossing process. The decrease in PSM value indicates a higher interaction with vehicles, which can lead to conflict with vehicles.

Figure 3.5 shows the process involved in PSM estimation. In the first step, a pedestrian arrives at an intersection crosswalk and observes for a gap. Once the pedestrian identifies a gap, he/she starts crossing at a time T_1. In the second step, the pedestrian reaches the second lane at the trajectory of an oncoming vehicle (conflicting point) at a time T_2. Once the pedestrian crosses the conflicting point of the vehicle, the vehicle arrives at the conflicting point at the time T_3. The difference between the time T_3 and T_2 is known as the PSM. A pedestrian might have multiple safety margins (one safety margin for each lane) until he/she reaches the median. The minimum of the three is the critical safety margin. For a two-way crosswalk, the PSM can be estimated using the same technique for the next crossing phase, i.e., median to end.

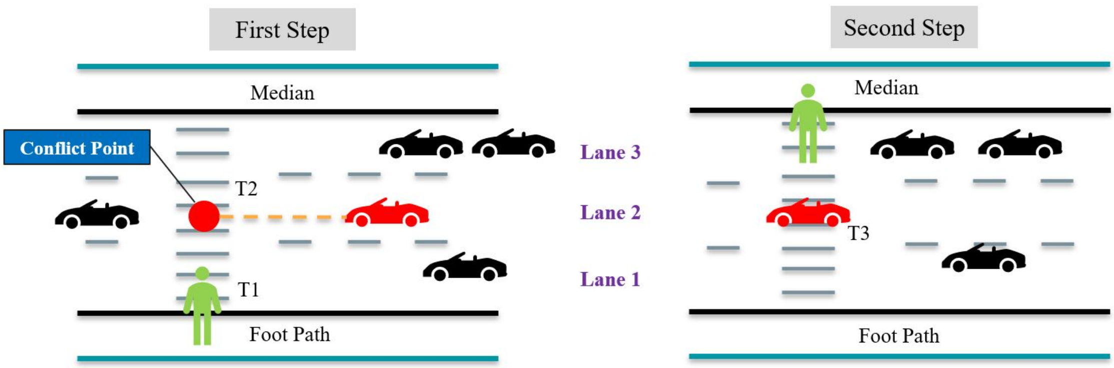


FIGURE 3.5: Pedestrian Safety Margin (PSM) estimation process.

In the current study, the PSM values are extracted at each stage of crossing, i.e., start to median and median to end for the two-stage crossing process. The least of all extracted values is considered for the PSM model development.

The majority of the studies in developed countries were conducted on well-designated signalised crosswalks. Furthermore, pedestrian-vehicle conflicts were few under uniform traffic flow conditions. On the contrary, in developing countries, heterogeneous traffic and non-lane discipline encourage a wide variety of road crossing behaviour, resulting in a higher rate of pedestrian interaction with the oncoming traffic.

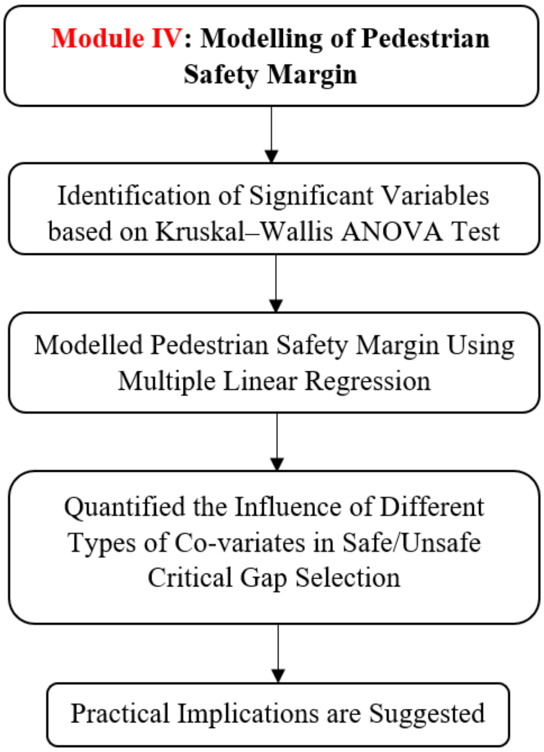


FIGURE 3.6: Module IV – Analysis framework for Pedestrian Safety Margin (PSM) modelling.

The PSM is a good measure of pedestrian-vehicle interaction at intersection crosswalks. Thus, the current study focused on the factors influencing the PSM with consideration of pedestrians at signalised intersection crosswalks. The complete methodology of the PSM study has been illustrated in Figure 3.6. The multiple Linear Regression (MLR) model is often used to model continuous data and explore the influence of various covariates. In the current study, the MLR modelling approach (for one-way crosswalks) and random intercept linear mixed-effects modelling approach (for two-way crosswalks) are utilised to understand the influence of different pedestrian and traffic level attributes on PSM.

3.3 Summary

The complete step-by-step structure of the adopted study methodology is presented in this chapter. An overall general methodology and study framework have been presented. Four separate frameworks were described in Modules I, II, III and IV. In Module I, the details of pedestrian signal violation behaviour are described. In Module II, pedestrian risky crossing behaviour in terms of distraction (digital and social) is explained. The various approach used for understanding pedestrian waiting behaviour is discussed in Module III. Module IV discusses the process for quantifying pedestrian-vehicle interaction using the Pedestrian Safety Margin approach. The location selection, data collection, extraction, and analysis procedure are explained in detail in the next chapter.

Chapter 4

Data Collection and Analysis

4.1 General

The previous chapter presented a brief overview of the methodology adopted to fulfil the study objectives, which include pedestrian signal violation behaviour, distracted road crossing, waiting behaviour and safety margin.

Accurate and reliable data is necessary to improve the safety, comfort, and convenience of pedestrians at signalised intersections. The data collection method requires a standard scientific procedure to ensure accuracy and reliability. Manual data collection is expensive, inaccurate, and time-consuming. To overcome these limitations, a video graphic data collection is recommended to collect required parameters at signalised intersections. This chapter describes the detailed process of data collection, extraction, and descriptive statistical analysis of extracted variables at signalised intersections.

4.2 Overview of Data Collection, Extraction, and Analysis Process

Identifying pedestrian crossing characteristics, traffic, and intersection geometry is essential for the crossing safety evaluation of pedestrians at intersection crosswalks. The overview of the data collection process (including site selection, data collection, extraction, and analysis) is presented in Figure 4.1 and explained in the following sections. The data was collected and extracted between 2018-2020.

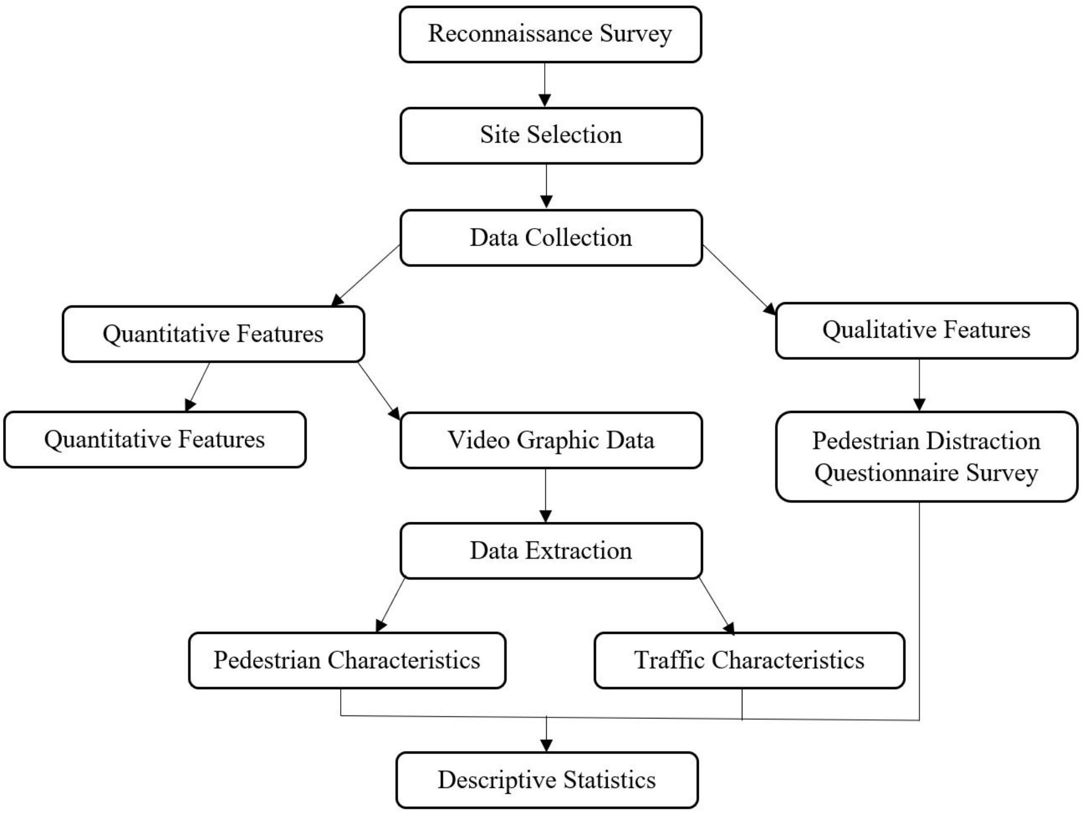


FIGURE 4.1: Step-by-step process of data collection and extraction.

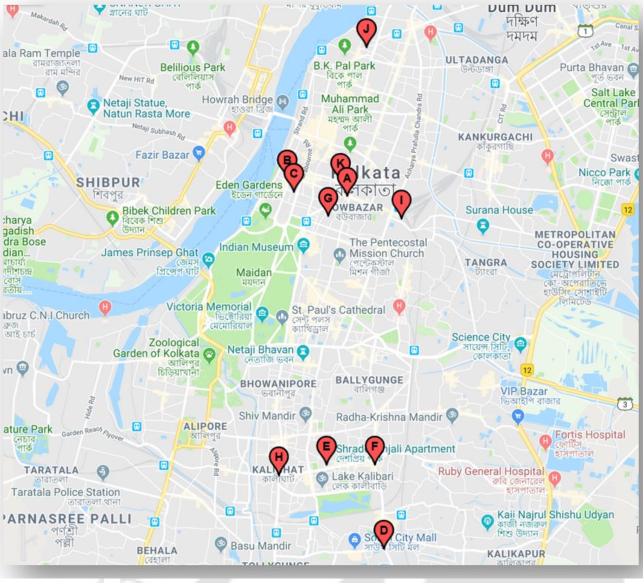
4.3 Identification of Survey Location

The inventory exercise is conducted for survey site selection. As per the literature section, West Bengal (Indian State) has the highest number of pedestrian fatalities reported in 2018. Thus, Kolkata (the capital of West Bengal) was selected for the current study observation. After obtaining ethical approval from traffic agencies, a team of researchers conducted a reconnaissance survey in Kolkata city in June 2018, and more than 25 intersections were visited to assess suitability. Different intersection and crosswalk characteristics were noted, such as location name, GPS coordinates, type of intersection, number of lanes, availability of pedestrian signal head and signal phase length. It was noted that in Kolkata, the intersections consisted of T, four and six-legged types. The intersection crosswalks are spread on three-lane undivided, four-lane divided and six-lane divided roads to serve various land use facilities on either side, such as building access points, shopping malls and transportation facilities (bus stops and metros). It is observed that most of these locations lack pedestrian signal heads, markings and signboards, which makes them unsafe for pedestrians. This research focuses on crosswalks with varied roadway characteristics with a wide range of pedestrian behaviour and mixed traffic conditions (two-wheelers, three-wheelers [auto ricksaws], cars and buses).

4.4 Site Selection

During the reconnaissance survey, ten-minute video footage of locations was collected to identify the best-suited sites for the study. The traffic and pedestrian flow and crossing characteristics have been identified. For the site selection, suitable intersections with minimum friction (roadside parking and traffic jam) are selected based on pedestrian and traffic flow. The site selection is limited to the signalised intersections in Kolkata city, with varied roadway characteristics from three-lane undivided one-way crosswalks to six-lane divided crosswalks. The variation in roadway characteristics is deemed essential to the range of traffic and pedestrian flow that directly affects pedestrian road-crossing behaviour. The intersection crosswalks were selected based on their moderate pedestrian and traffic volume and the availability of pedestrian signal heads. Based on the above criteria, 11 intersections were selected for data collection. The location name and number of crosswalks observed at each selected intersection are shown in Figure 4.2.

From all observed crosswalks, only three one-way and eight two-way crosswalks were selected (from 10 intersections) for the current study. The selected one-way intersections are BB Ganguly Street, General Post Office (GPO) and Dalhousie Square. All these intersections were situated in Central Business District (CBD) zone. The two-way intersections were Deshapriya Park, Jadavpur Thana More, Shobhabazar, Central, Chandni Chowk, Gariahat, General Post Office (near HSBC bank)) and Kalighat. The Central, Chandni Chowk, Gariahat, General Post Office and Shobhabazar intersections are located in the commercial area. Similarly, Deshapriya park and Kalighat are located in a mixed zone of commercial cum residential, while Jadavpur is located in a residential cum educational zone. The General Post Office intersection comprised of a one-way and a two-way crosswalk, and both were observed for the current study. All crosswalks were equipped with pedestrian exclusive signals with a countdown timer and had mixed traffic movement. None of the observed intersections had a pedestrian push button. The traffic comprised cars, two-wheelers, light commercial vehicles, and buses.



Mark	Intersection Location	Crosswalks Observed
A	BB Ganguly Street	2
B	GPO (BBD Bag)	2
C	Dalhousie Square	1
D	Jadavpur Thana	3
E	Deshapriya Park	3
F	Gariahat	3
G	Chandni Chowk	3
H	Kalighat	3
I	Sealdah	3
J	Shobhabazar	2
K	Central	3

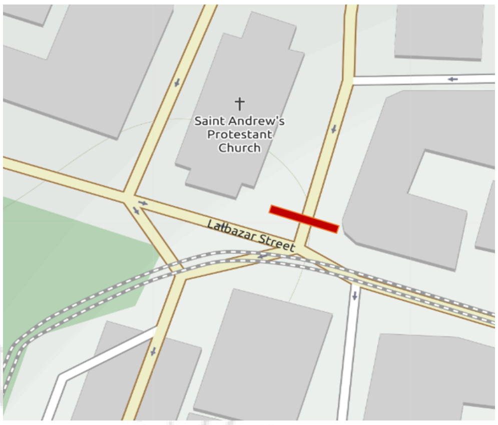
FIGURE 4.2: Kolkata study location site map.

4.5 Brief Details of the Selected Intersection Crosswalks

4.5.1 BBG Street Location

The site consists of one-way roads with approximately 1,100 vehicles per hour (average hourly volume for the observed period) and 1,120 pedestrians per hour (average hourly volume for the observed period) passing through the selected crosswalk. This intersection is located in the commercial CBD location. The selected crosswalk was adequately marked and equipped with pedestrian signals on both sides, as shown in Figure 4.3.

(A) Crosswalk view



(B) Location map

FIGURE 4.3: Location and overhead map view of BBG crosswalk.

4.5.2 General Post Office (GPO) Crosswalk Location

The site consists of one-way and two-way intersecting roadways, as shown in Figure 4.4. The selected one-way road serves approximately 665 vehicles per hour and 825 pedestrians per hour. The two-way road serves 3,300 vehicles per hour and 1,820 pedestrians per hour. This intersection is located in the commercial CBD location. The selected crosswalk was properly marked and equipped with pedestrian signals on both sides.

(A) Crosswalk view

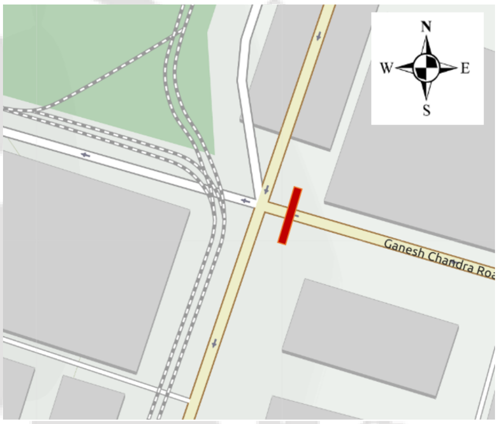
(B) Location map

FIGURE 4.4: Location and overhead map view of GPO crosswalk.

4.5.3 Dalhousie Square Location

The site consists of a one-way roadway with approximately 1,400 vehicles per hour (average hourly volume for the observed period) and 1,600 pedestrians per hour (average hourly volume for the observed period). The intersection is situated on the major road that connects the administrative zone of Kolkata, as shown in Figure 4.5. The crosswalk location is marked and equipped with signboards and pedestrian signals.

(A) Crosswalk view

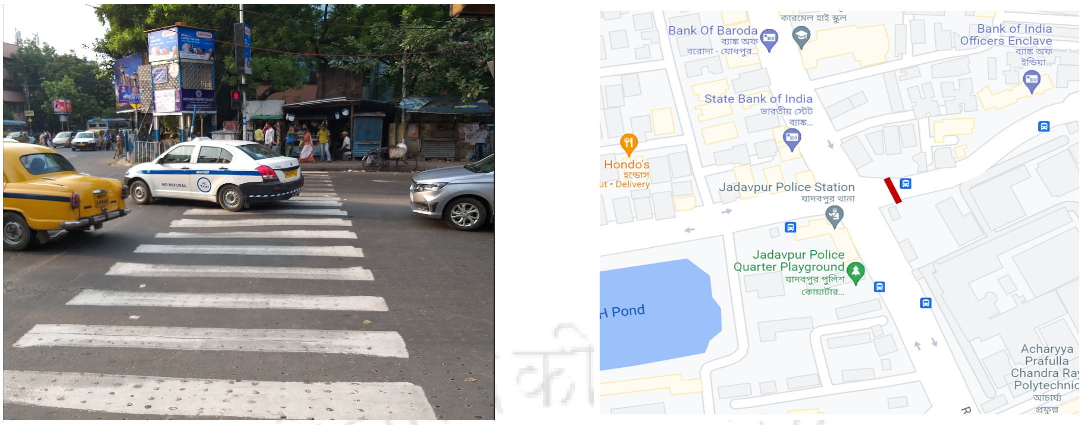


(B) Location map

FIGURE 4.5: Location and overhead map view of Dalhousie Square crosswalk.

4.5.4 Jadavpur Crosswalk Location

Jadavpur site comprised of roads with a four-lane divided roadway configuration. The selected road serves approximately 2,020 vehicles per hour and 360 pedestrians per hour. The intersection crosswalk provides access to the major residential and educational areas on both sides of the crosswalk (see Figure 4.6). At this site, major pedestrian movements occur during morning and evening peak hours. The selected crosswalk was marked and well-equipped with pedestrian signals.



(A) Crosswalk view (B) Location map

FIGURE 4.6: Location and overhead map view of Jadavpur crosswalk.

4.5.5 Deshapriya Park Crosswalk Location

Deshapriya Park site comprised of roads with a six-lane divided roadway configuration. The selected road serves approximately 2360 vehicles per hour and 530 pedestrians per hour. The intersection crosswalk serves as access to the residential and commercial zone on both sides of the crosswalk. The crosswalk is marked in Figure 4.7.

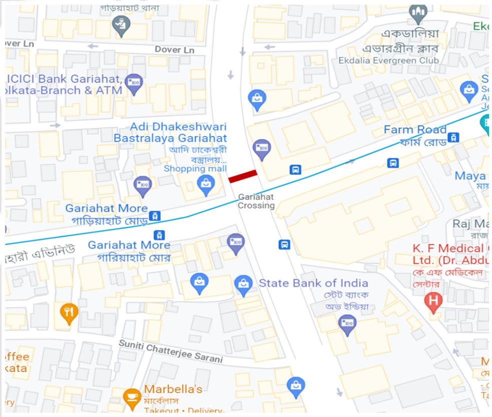
(A) Crosswalk view (B) Location map

FIGURE 4.7: Location and overhead map view of Deshapriya Park crosswalk.

4.5.6 Gariahat Market Crosswalk Location

Gariahat Market site comprised of roads with a six-lane divided roadway configuration. The selected road serves approximately 2,130 vehicles per hour and 2,020 pedestrians per hour. The intersection crosswalk serves as access to the major market areas on both sides of the crosswalk. This crosswalk is marked and well-equipped with pedestrian signals. A crosswalk view is shown in Figure 4.8.

(A) Crosswalk view

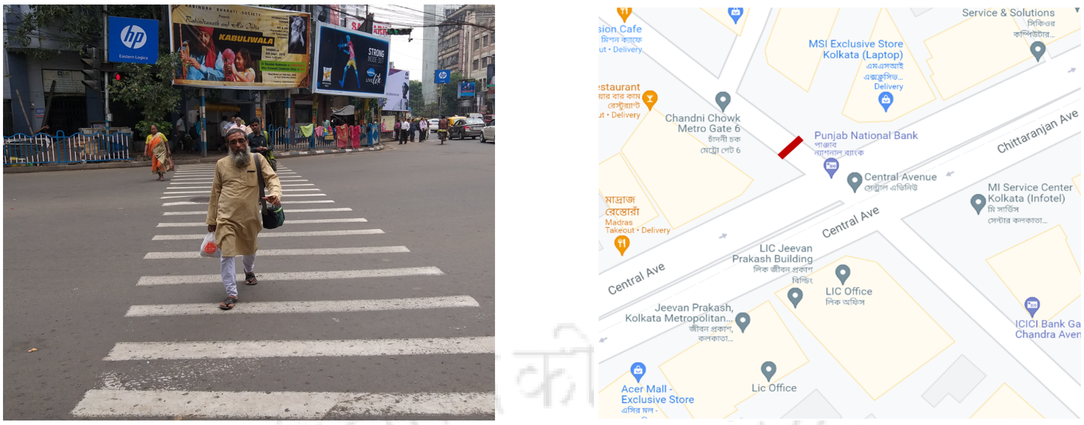


(B) Location map

FIGURE 4.8: Location and overhead map view of Gariahat crosswalk.

4.5.7 Chandni Chowk Crosswalk Location

Chadni Chowk site comprised of main roads with a six-lane divided roadway configuration. The selected road serves approximately 1,780 vehicles per hour and 410 pedestrians per hour. The intersection crosswalk serves as access to the major market area (commercial) on both sides of the crosswalk. The crosswalk is marked in Figure 4.9. Moreover, the selected crosswalk was properly marked and equipped with pedestrian signals on both sides.

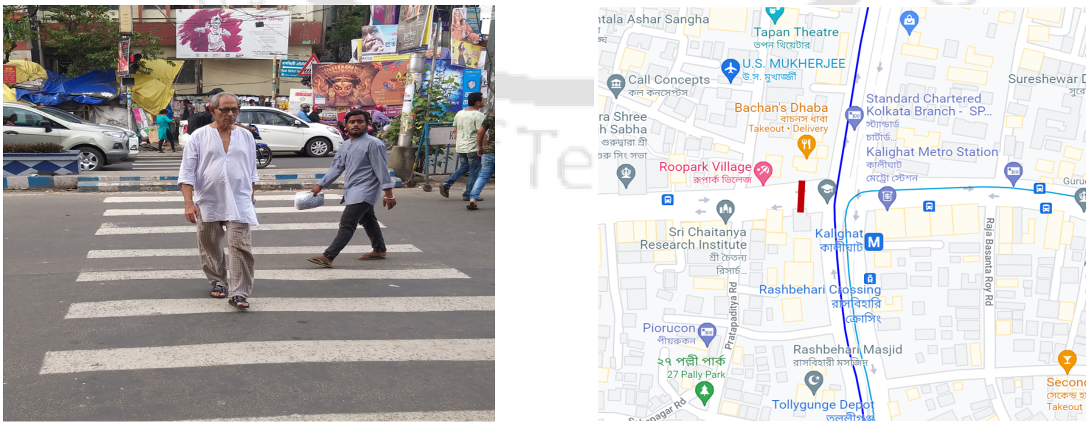


(A) Crosswalk view (B) Location map

FIGURE 4.9: Location and overhead map view of Chandni Chowk crosswalk.

4.5.8 Kalighat Crosswalk Location

Kalighat site comprised of roads with a four-lane divided roadway configuration. The selected road serves approximately 2,400 vehicles per hour and 460 pedestrians per hour. The intersection crosswalk provides access to the major residential and commercial areas on both sides of the crosswalk, as shown in Figure 4.10. This crosswalk also serves significant tourist attractions such as shrines and altars. The selected crosswalk was adequately marked, and it was well-equipped with pedestrian signals.

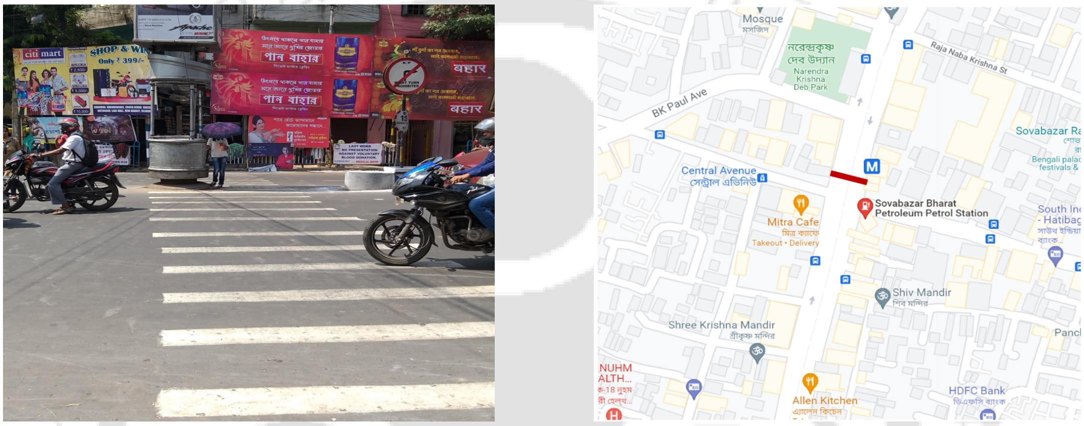


(A) Crosswalk view (B) Location map

FIGURE 4.10: Location and overhead map view of Kalighat crosswalk.

4.5.9 Shobhabazar Crosswalk Location

Shobhabazar site consists of roads with a six-lane divided roadway configuration. The selected road serves approximately 3,150 vehicles per hour and 860 pedestrians per hour. The intersection crosswalk provides access to the Shobhabazar Sutanuti metro station, as shown in Figure 4.11. The selected crosswalk was marked and well-equipped with pedestrian signals.



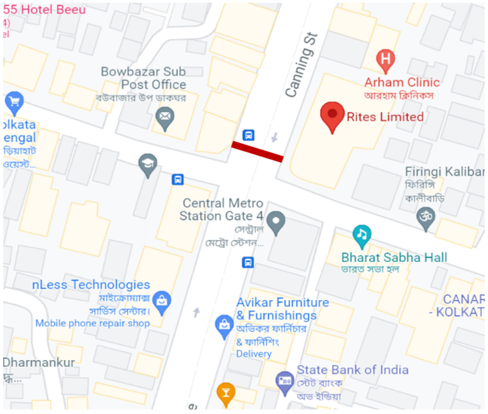
(A) Crosswalk view (B) Location map

FIGURE 4.11: Location and overhead map view of Shobhabazar crosswalk.

4.5.10 Central Crosswalk Location

Central site consists of roads with a six-lane divided roadway configuration. The selected road serves approximately 4,730 vehicles per hour and 650 pedestrians per hour. The intersection crosswalk provides access to the CBD location (commercial) on both sides of the crosswalk, as shown in Figure 4.12. The selected crosswalk was properly marked and equipped with pedestrian signals on both sides. The summary of all selected crosswalks is given in Table 4.1.

(A) Crosswalk view



(B) Location map

FIGURE 4.12: Location and overhead map view of Central crosswalk.

4.6 Data Collection

4.6.1 Pilot Survey

The pilot survey was conducted at each location to observe pedestrians as well as vehicular traffic flow and roadway characteristics. The collected pilot survey data was used to identify the appropriate intersections for the video graphic survey. Roadway characteristics (roadway width) were measured using a measuring wheel. Further, the pedestrian and traffic flow were measured in a laboratory setting from the recorded video. The video-based recording was selected as the primary method of data collection, which includes pedestrian and traffic characteristics.

4.6.2 Video Graphic Survey

After obtaining ethical approval from traffic agencies, a team of researchers recorded video footage using two high-definition fields mounted video cameras to gather pedestrian and traffic-related information inconspicuously. One camera was installed at a high elevation to get an overall view of the crosswalk. Another camera was fitted on a pole facing the crosswalk using a custom-built camera set-up installed at 12 feet above the ground to observe pedestrian demographics (gender and age), crossing characteristics, and vehicle movements accurately (see Figure 4.13). The video coverage included the pedestrian waiting area on the opposite side, the zebra crossing, and the pedestrian exclusive signal. This type of video-based data collection is advantageous over manual observations as video can be paused and rewound whenever needed, and many additional features can be later retrieved frame by frame in a laboratory setting. The footage was captured from 10 am to 2 pm (non-peak hours) for a single weekday (September 2018) during bright, dry weather conditions. In the current study, the peak hour time window (8:30 am – 10:00 am) was not considered for the study observation, as this time window exhibited high traffic and pedestrian volume, leading to frequent traffic jams at intersections. Additionally, Kolkata traffic police enforced special measures during peak hours at major intersections to reduce crashes, introducing interference in the independent observation. At busy intersections during the pedestrian red-light phase, traffic police pull out a rope to forcibly stop pedestrians from red-light jumping.

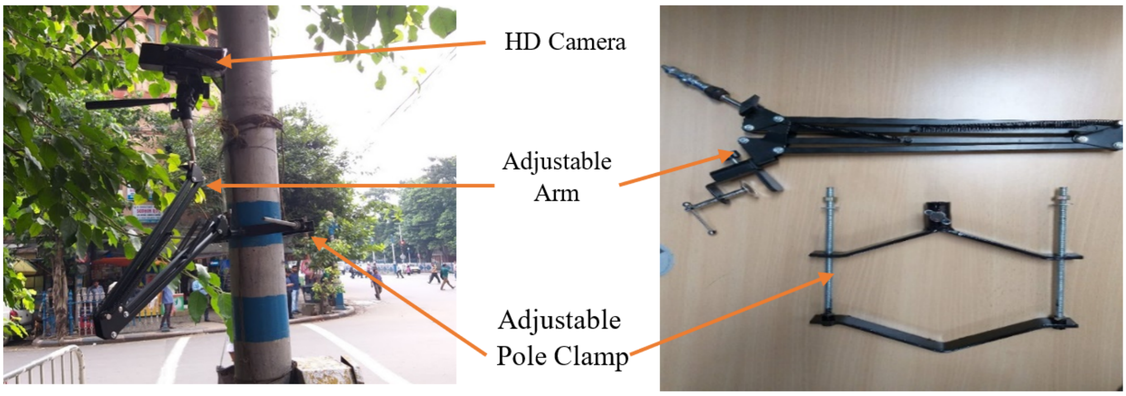


FIGURE 4.13: Pole camera setup for data collection.

TABLE 4.1: The roadway and signal characteristics of the selected crosswalks.

Crosswalk No.	Name of the intersection	Type of road	Land use type	Total number of lanes	Crosswalk dimention (m)	Mean cycle time (s)	Mean green time (s)	Mean red time (s)
1	Dalhousie	OW	C	Three	10.5 × 3	126 (32.7**)	82.7 (27.2)	42.9 (11.3)
2	BBG	OW	C	Three	10.6 × 3	141 (34.8)	85.3 (30.4)	55.2 (17)
3	GPO	OW	C	Three	11.7 × 3	90.9 (7.55)	48.5 (4.27)	42.4 (6.36)
4	GPO	TWD	C	Four	14 × 3	94.5 (20)	45.5 (11.7)	49 (10.8)
5	Central	TWD	C	Six	26 × 3	196.9 (55)	76.1 (32.6)	120.8 (38.6)
6	Chadni Chowk	TWD	C	Six	19 × 3	180 (15)	8.29 (1.24)	171.7 (15.2)
7	Deshapriya Park	TWD	C & R	Six	23 × 3	127.9 (28)	7.53 (1.36)	120.4 (27.6)
8	Gariahat	TWD	C	Four	17 × 3	190.2 (46)	14.6 (11.7)	175.6 (45.5)
9	Jadavpur	TWD	R & E	Four	17 × 3	136 (45)	9.73 (13.9)	126.3 (42.6)
10	Kalighat	TWD	C & R	Four	20 × 3	196.3 (61)	8.53 (2.79)	187.7 (59.2)
11	Shobhabazar	TWD	C	Six	19 × 3	103 (42)	47.3 (22.9)	55.7 (25.8)

Note:

OW: One-Way; TWD: Two-Way Divided; C: Commertial; R: Residential and E: Educational

**indicates Standard Deviation (SD) under each parenthesis.

4.6.3 Questionnaire Survey for Pedestrians' Distraction Exposure

4.6.3.1 Survey method and sampling

As vehicle drivers and pedestrians use a shared space for movement (especially at intersections); thus, similar to driver distraction (Horrey and Wickens, 2006; Pekker et al., 2011), the pedestrian could also engage in such distraction, which reduces crossing safety. The literature review highlighted that several studies had been carried out in developed nations displaying the potential impact of distracted road crossing behaviour on crossing safety. However, their conclusions were not validated with accident statistics due to the paucity of such data availability/inventory. Ralph and Girardeau (2020) raised similar concerns and pointed out that despite the lack of field evidence in terms of accidents data, scholars tend to describe distracted walking as a severe problem, which might misplace the key focus from the overall effort of improving safety for all road users. Therefore, in the absence of distraction-related accident records, a questionnaire survey of a representative sample could provide an overall accurate initial estimate of pedestrians' mobile phone engagement and its severity in daily commutes.

Thus, before collecting video-based evidence, a questionnaire survey was conducted across Kolkata to understand the impact of pedestrian distraction on daily commutes. A questionnaire was prepared, including two broad sections (A & B), where Section A covered demographic characteristics, including gender, age and profession. Likewise, Section B focused on understanding the type of phone a pedestrian use daily and the situation and motivation behind digital distraction. Additionally, pedestrians' past exposure to accidents or near misses due to digital distraction while walking/crossing a road was also included (refer Appendix A). Before conducting the survey, the designed questionnaire set was tested locally to optimise the number of questions and the time taken to complete the questionnaire. The questionnaire set (with closed-ended questions with multiple-choice options) was prepared in three languages (Bengali, Hindi, and English) to select participants from a wider spectrum of different cultural backgrounds. Seven undergraduates who could speak Bengali, Hindi, and English were trained to perform conversational-styled interviews. The conversational-style interview method was opted compared to the standardised interview method (constant worded) as this clarifies the intended meaning (Schober et al., 2004).

After finalising the survey locations, an interviewer-administered questionnaire survey of the pedestrians (those who are crossing through intersection crosswalks) was conducted in the intersection neighbourhood at each survey location (Figure 4.14). The survey took place during clear, dry weather conditions for a single weekday in each survey location between 11 am to 4 pm. The overall survey took eight days (for eight locations), and the survey dates were picked randomly over September month in 2018 (the start of winter). During the survey, participants were randomly selected (using a random sampling technique) and asked for participation, and those willing to undergo the interview process were finally interviewed. Due to the massive rush, the participation rate was low (i.e., 5-7% only). A ball-pen was given as a reward to all participants. In total, 570 pedestrians participated in the survey, and only 446 respondents answered all the questions thoroughly (completion rate: 78.2%). Afterwards, these 446 questionnaire samples were entered manually into an Excel sheet according to the final analysis requirement and used for the final exploratory data analysis.

(A) Enumerator asking questions to a female respondent

(B) Enumerator asking questions to a male respondent

FIGURE 4.14: Questionnaire field survey.

4.6.3.2 Sample size calculation

To get a reliable survey estimate, the survey sample size criteria must be satisfied. The sample size estimation depends on the population of interest (target population), error margin percentage, which tells about how much survey results reflect the views of the overall population and sampling confidence interval, which reveals the confidence that the population would select an answer within a specific range. The required sample size is estimated based on a modified version of the sample size calculation method reported by Taherdoost (2017), considering Kolkata's 15 million population size, 95% confidence interval, 5% margin of error and percentage occurrence of a state or condition (suggested value 0.5). The following equation 4.1 is used to estimate the survey sample size.

$$S = \frac{\frac{z^2 p(1-p)}{e^2}}{1+\left(\frac{z^2 p(1-p)}{e^2 N}\right)} \tag{4.1}$$

Where:

S = Required Sample size; N = Population size; e = Margin of error; z = Z-score (for 95% confidence interval, $z = 1.96$) and p = Percentage occurrence of a state or condition (suggested value 0.5).

The collected questionnaire survey sample size in the present study (446) is more than the required sample size (S: 385), which satisfied the sample size requirement.

4.7 Identification and Selection of Variables

The identification and selection of variables reflect the site selection process and the target outcomes. The selection of inappropriate variables for data collection would increase the data extraction time and lead to inappropriate results. The variable selection should be according to the study objectives. In this view, video-based data collection is a more appropriate data collection technique. To fulfil the objectives, the dependent variables were selected as signal violation (waited for green vs crossed in red), distracted road crossing or mobile phone use, waiting duration of the pedestrian in signal violation and safety margin value maintained by pedestrians while crossing the road. These variables can explain the pedestrian road crossing behaviour and safety while crossing the road.

The main focus of the current study is to understand the impact of various covariates on the selected dependent variables. The factors expected to affect the dependent variables are primarily categorised into three distinct categories, i.e., pedestrian crossing characteristics, traffic characteristics and intersection characteristics. Based on the study objectives and past studies, these variables were further categorised into eight distinct categories, which include pedestrian demographic, behavioural, social, violation, exposure, state of crossing, glance and critical events, and signal and traffic characteristics, presented in Figure 4.15. A detailed description of the variables is presented in Table 4.2.

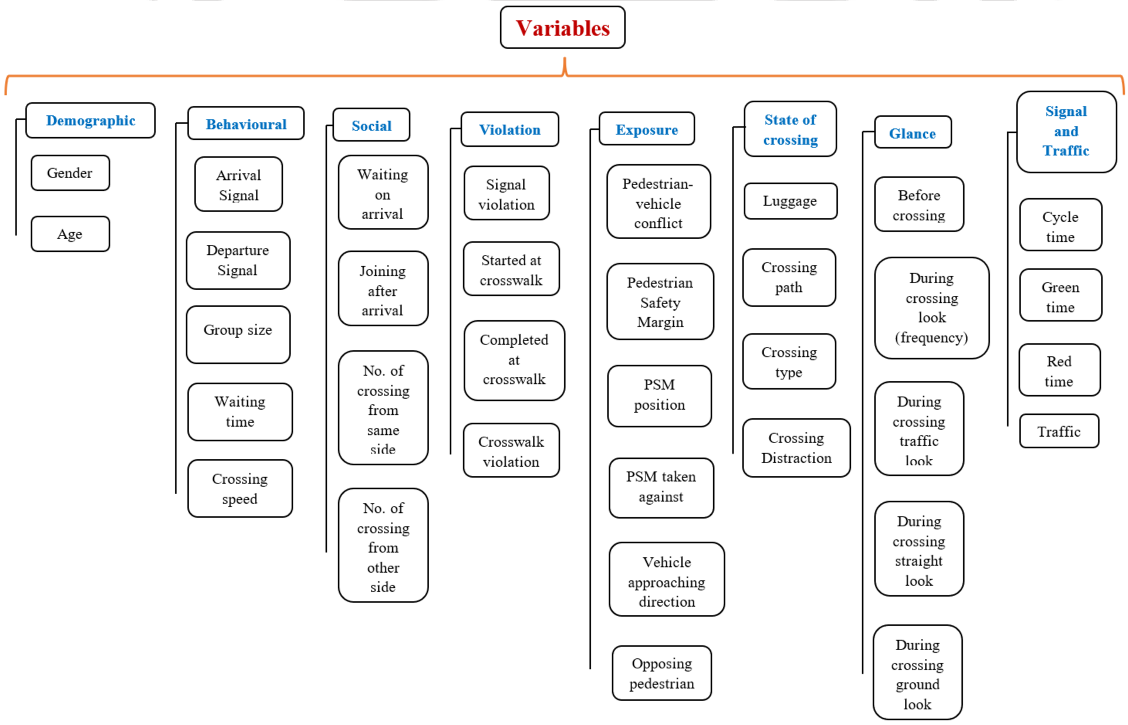


FIGURE 4.15: Factors and variables selected for the current study.

4.7.1 Demographic Characteristics

Demographic variables include gender and age of pedestrians. The age variable comprised of five distinct groups, i.e., $\leq$18 (young), 18-29 (young adults), 30-45 (adults), 46-60 (old adults) and $\geq$60 (old).

4.7.2 Pedestrian Behavioural Characteristics

Pedestrians' crossing behavioural characteristics included timestamps of different events like arrival and departure from the curb, and completion of individual pedestrian crossing (reaching the opposite end of the road). The crossing time (defined as the difference between crossing completion time and departure time) and the waiting time (till the green signal phase starts) were measured from the video. The width of the roadway was obtained from the geometric measurement at the site. The crossing times were used to estimate the crossing speed of pedestrians (roadway width divided by crossing time). Additionally, pedestrian social status, i.e., whether the pedestrian crossing alone or in a group (group/platoon size) was observed.

4.7.3 Social Information

The social information factors included the number of pedestrians waiting at the curb when a pedestrian arrived at the intersection in the red-light phase, illustrated in Figure 4.16 [Step 1]. The number of pedestrians joining upon arrival of the candidate pedestrian in the red-light phase, illustrated in Figure 4.16 [Step 2]. Additionally, the crossing of others in the red-light phase, i.e., the number of pedestrians (count) crossing from the same and opposite direction during the red-light phase [Figure 4.16, Step 3] was also considered.

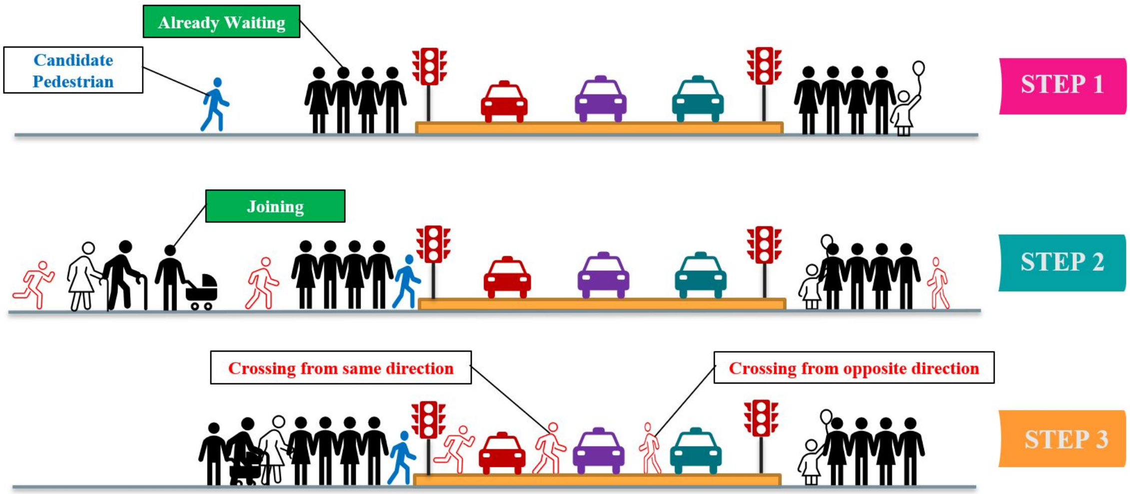


FIGURE 4.16: Social factors estimation process in the pedestrian red-light phase.

Note: It is assumed that in this red-light phase, the candidate pedestrian (blue colour) decided to wait with others for safe crossing.

4.7.4 Violation Characteristics

Pedestrians' lawbreaking/violation tendency was observed, including whether they went outside of the crosswalk (crosswalk violation) and violated the pedestrian signal (signal violation). It was assumed that if pedestrians took three or more steps outside the markings, they were considered to have not stayed within the markings. Similarly, pedestrians starting and ending the crossing process utilising marked crosswalks were also noted.

4.7.5 Exposure Characteristics

The exposure category includes variables that reflect the pedestrian and vehicle interactions. The interaction with vehicles has been identified based on whether the pedestrian experienced a conflict with the vehicle. Further, pedestrian risky crossing behaviour was captured using a dummy measure known as safety margin.

Additionally, vehicular type (for which safety margin were taken) and spacial factors (where the safety margin were taken) were noted down for each pedestrian safety margin. In addition, crossing friction (number of pedestrians crossing from the opposite direction) due to opposing pedestrian movement was also noted.

4.7.6 State of Crossing

The state of crossing includes (i) whether the pedestrian was carrying any sort of luggage, (ii) whether crossing along the crosswalk markings perpendicularly or diagonally (in an oblique path), (iii) crossing with a normal pace or hurried, and (iv) whether engaged in any sort of distraction.

Different types of distraction are defined as follows:

- **No distraction**: it represents a pedestrian who crossed without any distraction.
- **Mobile talking**: Holding a phone against the face and talking was categorised as talking on the phone.
- **Texting**: It is described as looking or interacting with the mobile screen.
- **Headphones**: If a pedestrian uses headphones (wired), it is categorised as headphones use; however, it is possible that a pedestrian could be using it for talking on the phone instead of listening to music. Additionally, no earbuds use was observed in the extracted observations, as it was not very much widespread during the survey year.
- **Holding a phone in hand**: Holding a phone in hand (not talking) while crossing was not a hazardous task. It was usually noticed that if a pedestrian was already distracted while approaching an intersection or waiting on

arrival, in such circumstances, when he/she decides to cross, he/she stops the interaction with the mobile device and holds it in hand. This type of communication break in the middle of continuous mobile engagement might increase the temptation of interacting with the device, even though the pedestrian was in the middle of the road crossing process. The degree of such temptation was impossible to measure from an observational study point of view; still, in the current study, this behaviour was included in the distraction category to understand if any behavioural difference exists with pedestrians with 'no distraction'.

- **Group talking/conversation**: If a pedestrian was crossing in a group and conversing with others, such distraction was categorised as 'group talking/-conversation'.
- **Eating/Drinking/Smoking**: Pedestrians observed eating, consuming beverages or smoking during crossing were added into this category.
- **Other**: The category of other distractions was utilised to capture distractions (looking inside the purse, reading and tending a child) that were not accounted for by any other categories.

4.7.7 Glance Behaviour Data

The glance behaviour is critical for gathering relevant information on the surroundings and traffic plying on the road. Two types of glance were noted down, which are before and during crossing glance. The during crossing glance was further divided into three categories (traffic glance, straight glance and ground glance) to gather more detailed information. In addition, during crossing glance frequency was also noted down.

4.7.8 Traffic and Signal Characteristics Data

Signal and traffic characteristics, which are part of non-social parameters, include signal cycle length, phase length (red-phase and green-phase), and traffic count per cycle on observed crosswalks.

4.8 Data Extraction

The video data were extracted with millisecond accuracy using 'MPC-HC' video player (30 frames per second) as shown in Figure 4.17. The traffic data and pedestrian behaviour observations were extracted by clicking the step forward option provided in the software for only one-directional movement (towards the camera). The data were extracted in a laboratory setting, and values were recorded in a pre-designed spreadsheet (refer Appendix B). Before the final video data extraction, inter-rater reliability among research associates (kappa) was established. A 96% inter-observer validity accuracy rate was found in labelling different types of crossing behaviours, age, distraction, and cautionary behaviour. To randomise the selection process, pedestrians were selected based on an odd sequence of arrival counts, such as first, third, fifth, seventh, and so on. Pedestrians who completed the whole crossing process and never meandered from the video frame were selected for feature coding. Pedestrian signal cycles were ignored from data analysis where vehicles had stopped (especially during the pedestrian green phase) on the pedestrian right of way, covering more than 50% width of the zebra crossing to wait for their green signal.

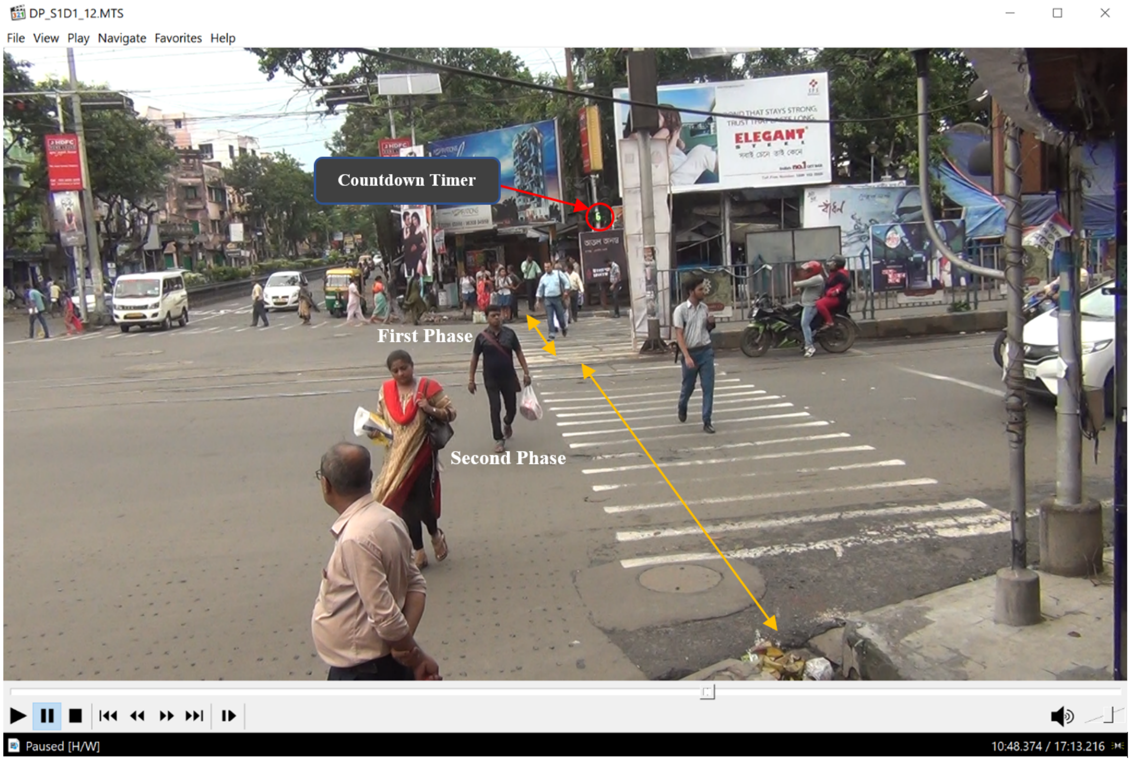


Figure 4.17: Data extraction overview using MPC-HC video editor.

TABLE 4.2: List of variable coded with description.

Characteristic	Variables	Type of variable	Description
Site	Site identifier	Categorical	1, 2, 3 and so on
Demographic	Gender	Categorical	Male/Female
	Estimated age (years)	Categorical	<18, 18-29, 30-45, 46-60, >60
Pedestrian behavioural characteristics	Signal on arrival	Categorical	Pedestrian signal on arrival at intersection
	Signal on departure	Categorical	Pedestrian signal on departure
	Signal on completion	Categorical	Pedestrian signal on crossing completion
	Group size	Categorical	Single and Two or more
	Waiting time (s)	Numeric	Time spent before crossing after arrival at crosswalk
	Crossing speed (m/s)	Numeric	The crossing distance (crosswalk length) divided by the crossing time
Social characteristics	No. of waiting upon arrival	Numeric	Number of pedestrians waiting upon arrival on red phase
	No. of joining upon arrival	Numeric	Number of pedestrians joining upon arrival on red phase

Table 4.2 continued from previous page

Characteristic	**Variables**	**Type of variable**	**Description**
	No. of crossing from the same direction	Numeric	Number of pedestrian crossings in the same direction in red phase
	No. of crossing from the opposite direction	Numeric	Number of pedestrian crossings from the opposite direction in red phase
Violation	Signal violation	Categorical	Yes/No
	Started on marked crosswalk	Categorical	Yes/No
	Completed on marked crosswalk	Categorical	Yes/No
	Crosswalk Violation	Categorical	Yes/No
Exposure	Conflict	Categorical	Whether a pedestrian experienced a conflict with traffic (Yes/No)
	Safety Margin	Numeric	Gap between the time a pedestrian cross before a vehicle and the time it arrives at the crossing point
	Safety Margin Position	Categorical	Initial lane, middle lane or end lane.

Table 4.2 continued from previous page

Characteristic	**Variables**	**Type of variable**	**Description**
	Vehicle type	Categorical	The type of vehicle yielded safety margin (car, two-wheeler, lcv and bus).
	Opposing pedestrians	Categorical	Number of pedestrians crossing from the opposite direction (none, one pedestrian, two or more) during crossing.
State of crossing	Carrying luggage	Categorical	Yes/No
	Crossing path	Categorical	Straight/Oblique
	Crossing type	Categorical	Normal/Running
	Crossing distraction	Categorical	No distraction, mobile talking, texting, headphones, group talking, eating/drinking/smoking, holding a phone and others.
Glance behaviour	Before crossing look at traffic	Categorical	Yes/No
	Before crossing look on both side	Categorical	Yes/No
	Before crossing look frequency	Numeric	Frequency of looking at traffic before starting to cross.
	Crossing look at traffic	Categorical	Yes/No

Table 4.2 continued from previous page

Characteristic	**Variables**	**Type of variable**	**Description**
	Crossing look at ground	Categorical	Yes/No
	Crossing look straight	Categorical	Yes/No
	Crossing look frequency	Numeric	Frequency of looking at traffic during crossing (head movement count).
Signal and traffic characteristics	Cycle time category	Categorical	0-100 s, >100-150 s and >150 s
	Green walk time category	Categorical	0-50 s, >50-100 s, >100 s
	Do not walk time category	Categorical	0-40 s and >40 s
	Traffic count per cycle	Numeric	Number of Vehicles plying on the street per cycle.

4.9 Data Analysis

In total, 25 hours of video data were extracted from the collected footage (9 hours for one-way and 16 hours for two-way crosswalks), resulting in 2360 observations for one-way and 2800 observations for two-way crosswalks. The descriptive statistics revealed that the majority of the observation consists of male pedestrians (one-way: 87.8% and two-way: 68.9%) and young adults age group between 30-60 years (presented in Table 4.3).

One primary reason for observing fewer samples of female gender groups in one-way crosswalk locations could be the land-use pattern of the sites. All the selected one-way sites were primarily located in the commercial region (Central Business District) of Kolkata city, where the Govt. and the corporate offices of multinational companies and banks are situated. Therefore, the trips to these CBD locations are majorly work-related and are dominated by males. In India, the gender gap in employment is vast and still increasing. As per the World Bank report 2020, women accounted for only 20.3% of the total workforce in India, which is significantly low compared to the world average (World Bank Group, 2020). Similar traits exist in Kolkata city too. In the current study, the lack of female observation could be due to a lack of female employees making work-related trips at the selected locations. Similar observations were also reported in other Indian cities as well, where male pedestrian observations were also dominant (Avinash et al., 2018; Marisamynathan and Vedagiri, 2018).

It is also observed that most pedestrians took caution (looked at traffic) before initiating the final crossing (one-way: 64.2% and two-way: 71.2%). The higher traffic glance at two-way crosswalks might be due to the presence of high traffic flow. Further, the primary segment of the observed pedestrians crossed the road

TABLE 4.3: Descriptive statistics of pedestrian observations.

Variable	One-way Crosswalks	Two-way Crosswalks		
Phase	*Overall*	*First Phase*	*Second Phase*	*Overall*
Gender				
Male	2073 (87.8%)	–	–	1930 (68.9%)
Female	287 (12.2%)	–	–	870 (31.1%)
Age				
<18	53 (2.2%)	–	–	196 (7.0%)
18-29	286 (12.1%)	–	–	720 (25.7%)
30-45	864 (36.3%)	–	–	1049 (37.5%)
46-60	856 (36.3%)	–	–	681 (24.3%)
>60	301 (12.8%)	–	–	154 (5.5%)
Group size				
Single	2014 (85.3%)	–	–	2190 (78.2%)
Two or more	346 (14.7%)	–	–	610 (21.8%)
Departure signal				
Walk	1749 (74.1%)	1211 (43.2%)	1253 (44.8%)	–
Don't walk	593 (25.1%)	1565 (55.9%)	1523 (54.4%)	–
Flashing	18 (0.8%)	24 (0.9%)	24 (0.9%)	–
Pedestrian crossing speed (m/s)				
Mean speed (SD)	1.30 (0.33)	1.36 (0.39)	1.30 (0.30)	1.30 (0.28)
15^{th} percentile speed	1.07	0.98	1.01	1.04
Critical safety margin (s)				
Mean safety margin (SD)	3.23 (1.89)	–	–	3.15 (1.64)
Crossing pace				
Normal	2186 (92.6%)	2537 (90.6%)	2624 (93.7%)	–
Hurried	174 (7.4%)	263 (9.4%)	176 (6.3%)	–
Crossing path				
Perpendicular	2094 (88.7%)	2354 (84.1%)	2532 (90.4%)	–
Oblique	266 (11.3%)	446 (15.9%)	268 (9.6%)	–
Crosswalk violation				
Yes	1032 (43.7%)	1017 (36.3%)	837 (29.9%)	–
No	1328 (56.3%)	1783 (63.7%)	1963 (70.1%)	–

TABLE 4.3: Continued from previous page.

Variable	**One-way Crosswalks**	**Two-way Crosswalks**		
Phase	*Overall*	*First Phase*	*Second Phase*	*Overall*
Social characteristics				
No. of pedestrians waiting upon arrival: Mean (SD)	4.54 (4.38)	–	–	–
No. of pedestrians joining upon arrival: Mean (SD)	6.47 (6.19)	–	–	–
No. of pedestrians crossing on red in the same direction: Mean (SD)	3.40 (3.52)	–	–	–
No. of pedestrians crossing on red in the opposite direction: Mean (SD)	2.76 (2.56)	–	–	–
Carrying luggage				
Yes	1265 (53.6%)	–	–	1684 (60.1%)
No	1095 (46.4%)	–	–	1116 (39.9%)
Distraction				
No distraction	1522 (64.5%)	1906 (68.1%)	1863 (66.5%)	–
Mobile talking	206 (8.7%))	139 (5.0%)	145 (5.2%)	–
Texting/watching screen	90 (3.8%)	38 (1.4%)	50 (1.8%)	–
Headphones	25 (1.1%)	49 (1.8%)	51 (1.8%)	–
Group talking	230 (9.8%)	271 (9.7%)	317 (11.3%)	–
Eating/drinking/smoking	71 (3%))	46 (1.6%)	51 (1.8%)	–
Holding a phone in hand	161 (6.8%))	235 (8.4%)	225 (8.0%)	–
Others	55 (2.3%))	116 (4.1%)	98 (3.5%)	–
Wait till green phase initiation (after red-phase arrival)				
Wait time: 0-20s	450 (49.1%)	462 (24.0%)	–	–
Wait time: 21-40s	353 (38.5%)	369 (19.2%)	–	–
Wait time: >40s	114 (12.4%)	1092 (56.8%)	–	–
Before crossing traffic look				
Yes	1515 (64.2%)	1993 (71.2%)	1542 (55.1%)	–
No	845 (35.8%)	807 (28.8%)	1258 (44.9%)	–
Traffic count				
Traffic count per cycle: Mean (SD)	29.2 (14.8)	–	–	110.18 (69.23)

alone (one-way: 85.3% and two-way: 78.2%), carrying some sort of luggage (one-way: 53.6% and two-way: 60.1%), and maintained an average speed of 1.30 m/s.

Out of 2360 observed pedestrians, 593 (25.1%) crossed the street, violating the pedestrian signal. For two-way crosswalks, the violation proportion for pedestrians was 55.9% in the first phase of crossing (start to median) and 54.4% in the second phase of crossing (median to end). The crossing characteristics variable revealed that the primary segment of the pedestrian crossed the street maintaining a normal pace following the zebra crossing.

The state of crossing distraction variable revealed that 34.5% of pedestrians showed some sort of distraction while crossing through one-way crosswalks. For two-way crosswalks, pedestrians showed higher distracted road crossing in the second phase of crossing (median to end: 33.5%) compared to the initial phase of crossing (start to median: 31.9%). The reason could be that in the median while searching for the crossing opportunity, pedestrians took out their phones to kill the undue halt time.

In the present study, the pedestrian-vehicle interaction is estimated using the Pedestrian Safety Margin (PSM) dummy measure. The descriptive statistics revealed that for two-way crosswalks, pedestrians took on an average smaller critical safety margin (mean: 3.15 s, SD: 1.64 s) compared to one-way crosswalks (mean: 3.23 s, SD: 1.89 s).

4.9.1 Distraction Questionnaire Survey Data Analysis

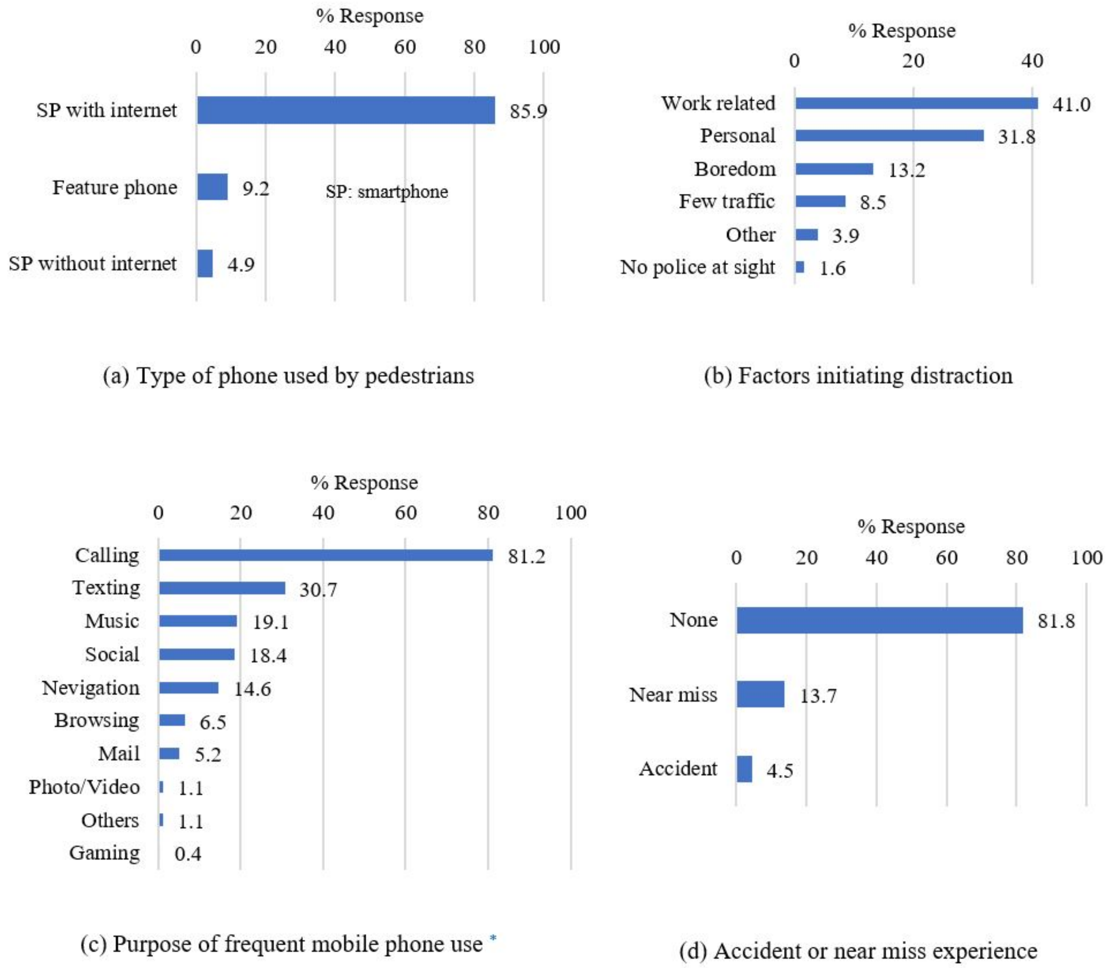
The questionnaire survey (Section A) analysis revealed that the participants were composed of 57.4% male and 42.6% of female (Table 4.4). The primary segment of pedestrians was young adults of 18-29 years (65.2%) and students (40.1%).

Table 4.4: Participants' demographic characteristics (sample size, n = 446).

Demographic variable and Frequency/Proportion (%)			
Gender			
Male	256 (57.4)	-	-
Female	190 (42.6)	-	-
Age		**Occupation**	
<18	19 (4.3)	Govt. Employee	35 (7.8)
18-29	291 (65.2)	Private Job	144 (32.3)
30-45	107 (24.0)	Self Employed	67 (15.0)
46-60	22 (4.9)	Student	179 (40.1)
>60	7 (1.6)	Others	21 (4.7)

The exploratory analysis of the mobile use questionnaire segment (Section B) revealed that the majority of the respondents use a smartphone (SP) with internet access, as presented in 4.18 (a). In response to the purpose of frequent phone usage, respondents reported that among all the reported mobile users, primary segment of the respondents was distracted from frequent mobile phone talking (81.2%); followed by texting (30.7%), listening to music (19.1%), social media usage (18.4%) and navigation (14.6%) activities (refer Figure 4.18 (c)). A primary segment of participants further reported that they were frequently involved in mobile phone use when it was work-related (41%), followed by personal reasons (31.8%) and out of boredom (13.2%); while standing alone or waiting for ride-hailing service (refer Figure 4.18 (b)).

The respondents were also asked about their experience with near-misses or accidents due to mobile use (digital distraction). The results are presented in Figure 4.18 (d). Out of all respondents (response: 446), 13.7% (61) stated that they had at least one near-miss in the past, and 4.5% (20) experienced at least one accident. The survey outcomes provided evidence that pedestrians do get involved in near-misses or accidents which may not be fatal but could lead to injuries and increase the overall risk of crossing.



(a) Type of phone used by pedestrians

(b) Factors initiating distraction

(c) Purpose of frequent mobile phone use *

(d) Accident or near miss experience

FIGURE 4.18: Mobile phone distraction survey response results (N = 446).

Note:
* Figure 1 (c) represents a question where a respondent could choose multiple options, for example a pedestrian could use mobile for talking as well as texting; thus, the aggregate proportion is not equal to 100%.

4.10 Summary

A detailed description of selected intersections is presented in this chapter, with the justification for selecting three one-way and eight two-way crosswalks. The details of crosswalk geometry, infrastructure, and sample photos with descriptions are presented in the current chapter. The procedure of video-graphic and questionnaire survey data collection, data extraction are explained in detail. Finally, descriptive statistics of pedestrian characteristics of 2360 pedestrians for one-way crosswalks and 2800 pedestrians for two-way crosswalks were described. Descriptive statistics with data analysis was done to understand pedestrian crossing behaviour at intersection crosswalks. The pedestrian risky crossing behaviour modelling is described in the subsequent chapters.

Chapter 5

Social and Non-Social Factors Influencing Signal Violation Behaviour

5.1 General

The previous chapter discussed intersection selection, data collection, coding and analysis. The chapter also discussed the variables extracted and coded for the current study. The current chapter discusses the statistical tests for selecting significant model-development variables from the collected data. The primary objective is to identify significant variables for model development. Statistical programming languages R (version 4.1.1) and Stata 17 have been utilised for preliminary analysis purposes.

5.2 Outlying Objectives

Based on the literature gaps discussed in Section 2.6, in this chapter, the following three hypotheses have been investigated in the context of low and middle-income countries:

- Hypothesis 1 (H_1): the crossing decision of an oncoming pedestrian depends on whether the pedestrian is waiting alone at the curb or waiting with others (other pedestrians are already waiting before the candidate pedestrian arrived or joined later) standing nearby.
- Hypothesis 2 (H_2): the likelihood of crossing in the red-light phase increases when one finds others (a significant number of pedestrians) crossing successfully in the red-light phase from the same or opposite direction.
- Hypothesis 3 (H_3): as the waiting time for safe crossing (time until the green-signal phase start) increases, the likelihood of signal violation also increases.

The first and second hypotheses have been evaluated in past studies (Rosenbloom, 2009; Faria et al., 2010); however, sample sizes were limited and predominantly conducted in developed countries. The current study investigates these hypotheses (H_1 and H_2) at multiple intersections with a larger sample size. Under the second hypothesis evaluation, the count of pedestrians crossing against the red-light phase is considered (as against the binary count considered in the previous study, (Rosenbloom, 2009)), which may provide a better explanation of signal violation behaviour. The impact of non-social features considered in hypothesis three (H_3) has been attempted in some past studies. However, the current study considered the impact of time-dependent signals installed at selected intersections, where traffic regulators update signal phase duration based on traffic volume throughout the day by prioritising vehicular traffic. Therefore, the current study tried to evaluate all three hypotheses (Hypothesis 1 to 3) formulated above from the data collected on pedestrian crossing behaviour at signalised intersection crosswalks located in Kolkata, India.

5.3 Selection of Significant Variables for Signal Violation Modelling

The selection of significant contributing variables is essential to develop better predictive models. The most frequently used tools are Pearson correlation (used to check the multicollinearity of continuous variables), contingency table (to assess the association/difference between categorical variables), t-test (to assess the difference in group mean) and Analysis of Variance (ANOVA).

5.3.1 Statistical Test for Signal Violation Model Variables

It is important to select the significant contributing variables for the signal violation model development. In general, increasing the number of variables improves the model fit but also increases the model complexity, hence often overfits the model. The trained model often performs well on training data but fails while generalising on real-world test data. Hence, it is crucial to calibrate the model with appropriate significant variables. The tests are selected based on the nature of the outcome (continuous/categorical) as well as explanatory/independent variables. The probability of signal violation (violated/ not violated) is represented using binary discrete outcomes. Thus, cross-tabulation test (viz., binary type independent variables) and t-test (viz., continuous type independent variables) were selected to test the significance of the selected variables (viz., gender, age, social variables etc.) over the dependent variable (violated/ not violated). The cross-tabulation method uses frequency count for making comparison between dependent and independent categorical variables. The most common test is Pearson's chi-square test of independence. It is a non-parametric hypothesis test that measures the association and differences between observed and expected values. Pearson's Chi-Square hypothesis can be tested as:

The null hypothesis (H_0): The two categorical variables are independent, and there is no association between them (the distribution of one variable is not influenced by the variation in the other variable).

The alternative hypothesis (H_1): The variables are not independent, and there is an association between them (distribution of one variable influenced by another variable).

The cross-tabulation is applied to the signal violation data (918 data points). A significance level of 90% or above has been used for testing the hypothesis. The cross-tabulation results are shown in the Table 5.1. The table results revealed that no significant difference was observed between signal violation and gender, i.e., between male and female pedestrians. The Pearson Chi-Square value has been found as $\chi^2 = 0.119$ (p = 0.729). In the case of age and 'group size' variables, a significant difference has been observed at 90% and 95% CI. Based on the Chi-Square test results, variables such as 'carrying a luggage' (χ^2 = 20.76, p = 0.150) and 'crosswalk violation' (χ^2 = 0.004, p = 0.946) revealed that there is no significant effect on pedestrian signal violation behaviour. Pedestrian signal violation probability significantly differs for pedestrians who crossed following a straight path compared to an oblique path (χ^2 = 2.908, p = 0.088). The test results also revealed that the influence of 'waiting time till green signal phase' (0-20s, 21-40s and >40s) on signal violation was significantly different for different waiting groups (χ^2 = 41.412, p = <0.001).

TABLE 5.1: The cross-tabulation test results for pedestrian signal violation

Variable	Pearson Chi-Square	DF	P-value
Gender	0.119	1	0.729
Age	8.899	4	0.064*
Group size	12.153	1	<0.001**
Luggage	2.076	1	0.150
Crossing path (straight/oblique)	2.908	1	0.088*
Crossing pace (normal/hurried)	13.859	1	<0.001**
Crosswalk violation	0.004	1	0.946
Crossing speed	17.486	4	0.002**
Waiting time till green phase initiation	41.412	2	<0.001**
Cycle time category	10.218	2	0.006*
Red time category	7.345	1	0.007*

Note:

DF: Degree of Freedom

** Denotes variable significance at 95% confidence level;
* Denotes variable significance at 90% confidence level.

Pedestrian 'crossing speed' ($\leq$ 1m/s, >1-1.2m/s, >1.2-1.4m/s, >1.4-1.5m/s, >1.5m/s), 'pace' of crossing (normal/ hurried), cycle time (0-100s, 101-150s and >150s) and red time (0-40s and >40s) also have significant influence on signal violation behaviour.

Further, the significance of various continuous variables is assessed using an independent sample t-test. The independent sample t-test (Levene's test) revealed that social factors and 'traffic count per cycle' are significantly different based on their mean and standard deviation of the two observed groups (viz., violated and not violated). The test results are presented in Table 5.2, indicating equal variances have not been assumed in pedestrian signal violation behaviour.

From the statistical test (viz., cross-tabulation and t-test), the variables such as age ($\leq$ 18, 18-29, 30-45, 46-60 and >60), group size (single and two or more), crossing

Table 5.2: The independent t-test (Levene's) for equality of variance results for pedestrian signal violation

Variable	Mean	SD	P-value
No. of pedestrian waiting upon arrival	4.450	4.380	<0.001**
No. of pedestrian joining upon arrival	6.470	6.190	0.034**
No. of pedestrian crossing from same direction	3.400	3.520	<0.001**
No. of pedestrian crossing from opposite direction	2.760	2.560	0.013**
No. of traffic plying per cycle	110.180	69.230	0.001**

Note:

** Denotes variable significance at 95% confidence level;
* Denotes variable significance at 90% confidence level.

path (straight/oblique), crossing pace (normal/hurried), crossing speed ($\leq$ 1m/s, >1-1.2m/s, >1.2-1.4m/s, >1.4-1.5m/s, >1.5m/s), waiting time till green signal phase initiation (0-20s, 21-40s and >40s), cycle length (0-100s, 101-150s and >150s), number of pedestrian waiting upon arrival of the candidate pedestrian, number of pedestrians joined later, number of neighbours crossing in red-signal phase, number of crossing from the opposite direction in red-signal phase and traffic count per cycle are selected for the pedestrian signal violation model development.

5.4 Association Among Variables

In addition to variable selection, it is a common practice to estimate multicollinearity among predictors using a pairwise correlation. This is done to eliminate predictors having a correlation value beyond a threshold limit (rule of thumb, $r \geq 0.4$). The pair-wise correlation among numeric variables (i.e., number of pedestrians waiting upon arrival, number of pedestrians joining upon arrival, number of pedestrians crossing from the same direction, number of pedestrians crossing from opposite direction and number of vehicles plying per cycle) revealed that Pearson's correlation (r) is $\leq$0.30.

In the present study, most of the variables are of categorical type, thus instead of estimating pair-wise Pearson correlation coefficients, Theil's U (Theil, 1958, 1966) association statistic was estimated using equation 5.1.

$$U(X|Y) = \frac{S(X) - S(X|Y)}{S(X)} \tag{5.1}$$

Where $S(X)$ is the entropy of X and $S(X|Y)$ is the conditional entropy of X given Y.

Theil's U gives an association strength for categorical variable pairs ranging from 0 to 1, where 0 indicates no association (i.e., y provides no information about x) and 1 indicates perfect association (i.e., y provides complete information about x). It is an asymmetric, [U$(x, y) \neq$ U(y, x), where U is Theil's U] measure of association. The pair-wise Theil's U association were estimated and reported in Figure 5.1.

The pair-wise values in the plot revealed that the association strength among categorical variables is weak (Theil's U: $\leq$0.20), indicating that the selected variables are independent and can be used in the model estimation process.

5.5 Model Formulation for Signal Violation Behaviour

The objective of the current study is to understand the influence of non-social and social features on pedestrian signal violation behaviour. As the signal violation is binary (violated/not-violated) in nature, to model a discrete binary outcome, binary logistic regression was used (Ciaburro, 2018; Washington et al., 2011). The equation used in a binary logistic regression model to estimate the probability of

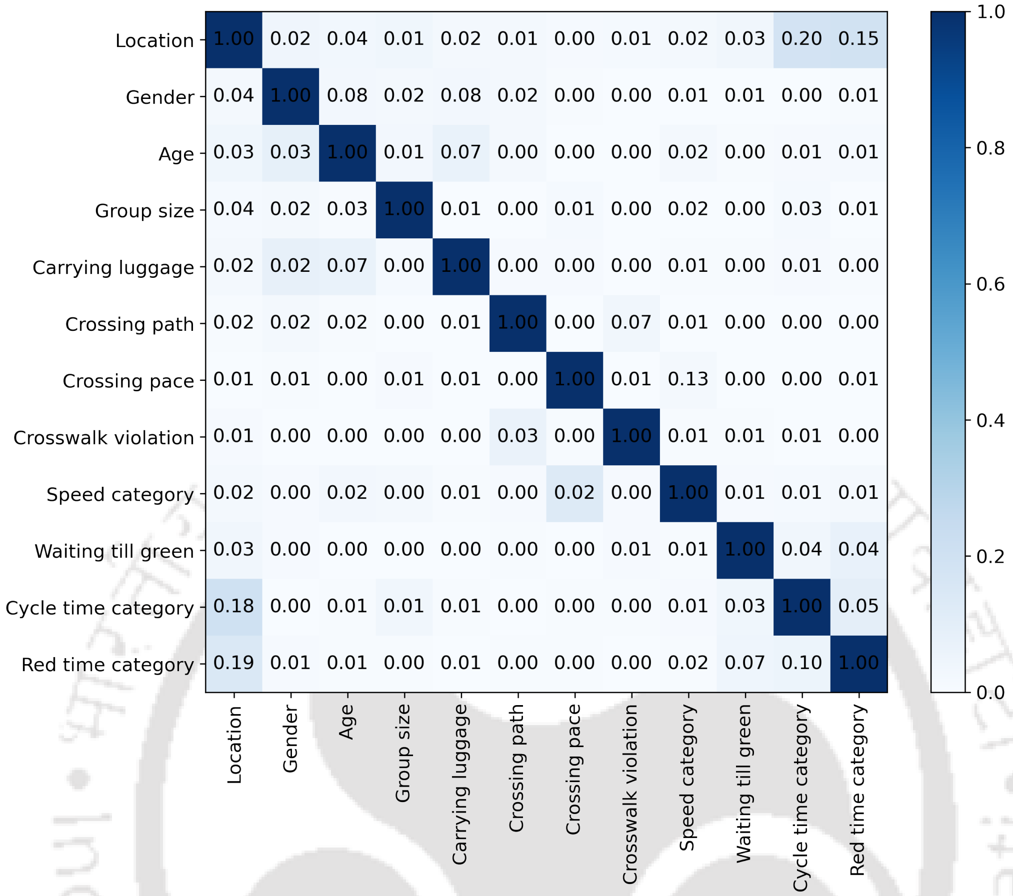


FIGURE 5.1: Theil's U pair-wise association plot.

signal violation is shown in equation 5.2:

$$P_i = \frac{\exp^{(\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i} \ldots +\beta_k X_{k,i})}}{1 \ + \ \exp^{(\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i} \ldots +\beta_k X_{k,i})}} \tag{5.2}$$

where P_i is the probability of pedestrian i committing a signal violation. β_0 is the model constant, and $\beta_1, \ldots, \beta_k$ are coefficients estimated by maximum-likelihood from the corresponding explanatory variables $X_1 \ldots X_k$. The logistic regression was estimated using the statistical programming language R (version 4.1.1), and Average Marginal Effects (AME) were calculated at every observed value and averaged using R's "margin" library (Leeper, 2018a).

5.6 Goodness of Fit Measure

After identifying the relevant variables, the developed model accuracy is tested with the help of suitable goodness of fit measure. The goodness of fit statistics, such as the rho-square and chi-square tests, are often used to test the calibrated models. Along with these measures, a significance test must be conducted for the covariates to decide whether these variables are statistically significant or not to be included in the final model. Generally, it is done using t-test statistics and p-value. In addition, often Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) are also reported as model fit statistics.

5.6.1 Rho-square Statistic

The goodness of fit statistics are used to evaluate how well the model fits the data, and usually, log-likelihood is used to measure the goodness of fit statistics. For the binary logistic regression model, rho-square is often used as a goodness of fit measure. The index range between 0 and 1. There are various rho-square statistics proposed, such as McFadden, Cox and Snell and Nagelkerke. McFadden Pseudo rho-square is the most frequently reported goodness of fit metric. The rho-square statistic compare model performance with all fitted parameters and the model with zero value of all the parameters. The rho-square statistic is computed using the following equation:

$$Rho - square = 1 - \frac{LL(\beta_0)}{LL(\hat{\beta})} \tag{5.3}$$

Where: $LL(\hat{\beta})$ = log-likelihood function with estimated parameters, $LL(\beta_0)$ = log-likelihood function with all parameters as zero.

5.6.2 The Likelihood Ratio Test (LRT)

The likelihood ratio test compares the goodness of fit of two statistical models. The LRT compares two hierarchically nested models to determine whether or not adding more parameters (adding complexity) makes the model significantly more accurate. The test statistic is estimated by taking the ratio of likelihoods of two models. It is computed using the following equation:

$$X_L = -2\left(LL(\beta_0) + LL(\hat{\beta})\right) \tag{5.4}$$

Where $LL(\beta_0)$ is the likelihood of a null model when all parameters are set to zero and $LL(\hat{\beta})$ is the log-likelihood of the model with all fitted parameters on convergence. The log-likelihood ratio test statistics and p-value of statistical significance of 0.05 ensure that at least one significant parameter exists in the model.

5.6.3 Akaike Information Criteria (AIC)

The AIC metric is used to compare models. The AIC is proposed by Akaike (1974). It is a measure of the goodness of fit of a statistical model. The AIC is calculated as:

$$AIC = -2\ln L + 2k \tag{5.5}$$

Where $\ln L$ is the maximised log-likelihood of the model, and k is the number of parameters estimated. The model with the smallest AIC value is considered the best-fitted model.

5.6.4 Bayesian Information Criteria (BIC)

BIC is a criterion functioning similar to AIC (Schwarz, 1978). The BIC is calculated as:

$$BIC = -2 \ln L + k \ln N \tag{5.6}$$

where $\ln L$ is the maximised log-likelihood of the model, k is the number of parameters estimated, and N is the sample size. A lower BIC value represents a better model fit.

5.7 Model Estimation

The binary logistic regression model was fitted using R statistical programming language with the significant variables obtained from Pearson's chi-square test of independence and Levene's t-test. Further, to improve the model fit (reduce model complexity), a backward step-wise elimination process was utilised. In the step-wise elimination process, first, the model was fitted with all variables obtained from the variable selection process. Subsequently, certain model variables have been removed based on the judgement of whether their removal caused a significant change in other variables or aggregate model and model fit statistics (AIC). In the step-wise fitting process in addition to significant variables (obtained from the variable selection process), gender, age and site variables were also included to check for improvements in the aggregate model. It was observed that adding site and gender variables improves the model fit and improve other variables' significance. Thus, they were kept in the final model, while age did not show any influence on the model fit, thus excluded.

5.8 Model Results

Pedestrians crossing in the red-signal phase or waiting for the green phase were analysed using a binary logistic regression model. Signal status (red/green) during crossing was selected as the dependent variable. The model is fitted with all red-light phase arriving observations (sample size of 918). The model outcomes indicated a wide range of variables capturing the effects of demographic, crossing characteristics, social characteristics, violation (crosswalk), and signal and traffic characteristics influencing the pedestrian signal violation at signalised junctions in Kolkata, as reported in Table 5.3. A significant number of variables included in the model have a p-value smaller than 0.10. The McFadden Pseudo R-squared value (0.24) and the log-likelihood ratio test ($p<0.05$) indicate an overall good model fit. The estimated AIC and BIC values are 975.732 and 1067.334, respectively. In the present model, a positive sign of a coefficient of an independent variable indicates that the factor increases the likelihood of signal violation. In contrast, the negative sign indicates the factor reduces the likelihood of signal violation.

The gender variable is observed to be insignificant in predicting pedestrian signal violation behaviour. Results highlighted that pedestrians exhibited higher crossing speed during the signal violation. The number of pedestrians who had been already waiting before the pedestrian's arrival, as well as the number of pedestrians who joined after the subject pedestrian's arrival, negatively influenced their signal violation decision, which confirmed the "Hypothesis 1" depicting that the crossing decision of oncoming pedestrians depends on whether they are waiting alone at the curb or waiting with others. Further, social cues obtained from red-light violators, who crossed from the same direction, positively influenced signal violation behaviour, hence, confirming "Hypothesis 2".

Table 5.3: Signal violation logistic regression model summary

Features	**Variable name**	**Coefficient (β)**	**Standard error**	**P-value**	**AME**
	Constant	1.593	0.571	0.005	-
Site	Site 1 (Dalhousie) **	0.508	0.247	0.040	0.076
	Site 3 (GPO)	-0.550	0.414	0.184	-0.091
Demographic	Gender: Female	0.011	0.265	0.968	0.002
Crossing characteristics	Crossing path: oblique **	0.664	0.230	0.027	0.101
	Crossing pace: hurried **	1.438	0.413	<0.001	0.200
	Crossing speed: >1-1.2 (m/s)	0.137	0.247	0.580	0.023
	Crossing speed: >1.2-1.4 (m/s) *	0.455	0.263	0.084	0.076
	Crossing speed: >1.4-1.5 (m/s) **	0.799	0.383	0.037	0.128
	Crossing speed: >1.5 (m/s) **	0.617	0.308	0.045	0.101
	Waiting time till green phase: 21-40 s **	1.095	0.199	<0.001	0.183
	Waiting time till green phase: >40 s **	1.872	0.371	<0.001	0.284
Social Characteristics	No. of pedestrian waiting upon arrival **	-0.078	0.022	<0.001	-0.013
	No. of pedestrian joining after arrival **	-0.096	0.015	<0.001	-0.015
	No. of pedestrian crossing in the same direction during red light phase **	0.100	0.026	<0.001	0.016
Signal and Traffic Characteristics	Cycle time: >100-150 s **	1.125	0.307	<0.001	0.187
	Cycle time: >150 s **	3.103	0.463	<0.001	0.404
	Through traffic per cycle **	-0.085	0.012	<0.001	-0.014

Model fit statistics
McFadden (Pseudo R-squared) = 0.244, AIC = 975.732 and BIC = 1067.334

Likelihood ratio test
Df.diff = -18, LogLik.diff = -128.192, χ^2 = 56.38 and P-value = <0.001

Reference levels
site = site2; gender: male; crossing path = straight; crossing type = normal; crossing speed ≤1.0 m/s; Waiting time till green light = 0-20s; cycle time category 0-100s

Note:

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

Regarding time-dependent signals, 'waiting time till green phase initiation' and signal cycle length significantly influenced the signal violation tendency. The positive significant coefficients of the "waiting time till green initiation" variable confirmed "Hypothesis 3". Traffic volume variable negatively influenced the signal violation behaviour. The model outcomes, along with their detailed discussion, are presented in the subsequent sections.

5.8.1 Demographic Characteristics

Similar to the observations made by Dommes et al. (2015) and Ren et al. (2011), in the present study, the gender of pedestrians did not emerge as an essential factor in signal violation. However, other observational studies investigating signal violation behaviour found a significant impact of gender during signal violation (Brosseau et al., 2013; Rosenbloom, 2009).

5.8.2 Crossing Characteristics

Model estimates presented in Table 5.3 revealed that pedestrians who can maintain a crossing speed between 1.4-1.5 m/s and above 1.5 m/s are 12.8% and 10.1% more likely to violate signals than those with a crossing speed of $\leq$1 m/s. Li and Fernie (2010) also reported similar findings in their study performed in Toronto (Canada). The findings revealed that non-complying pedestrians walk relatively faster (1.69 m/s) than compliant pedestrians (1.29 m/s). Pedestrians exhibiting haste (rushed) behaviour tend to violate the signal and take smaller gaps in the traffic stream to complete their manoeuvre quickly. This behaviour can also be corroborated by the crossing pace variable (normal/hurried) reported in the model estimate table (Table 5.3). The variable estimate of haste (hurried) behaviour showed that when pedestrians are in a hurry, they are 20% more likely to violate the signal than pedestrians crossing normally.

5.8.3 The Presence of Pedestrians Waiting/ Joining Afterward

The negative coefficients of the model estimate confirmed "Hypothesis 1", stating that an oncoming pedestrian is more likely to wait for the green phase if a considerable number of pedestrians are already waiting or joining at the curb afterwards. Figure 5.2 illustrates that as more people wait at the curb for the green light, the signal violation likelihood of an oncoming pedestrian decreases. Similarly, over the waiting period, as more pedestrians start joining the candidate pedestrian, the likelihood of signal violation decreases as illustrated in Figure 5.3. Different explanations can be given for such behaviour. The theory of social control explained the mechanism behind obedient behaviour as the motivation to be rewarded just for being conformist (Hirschi, 1969). Normal individuals have inner controllers that prevent them from breaking the law and encourage them to behave in a normative fashion.

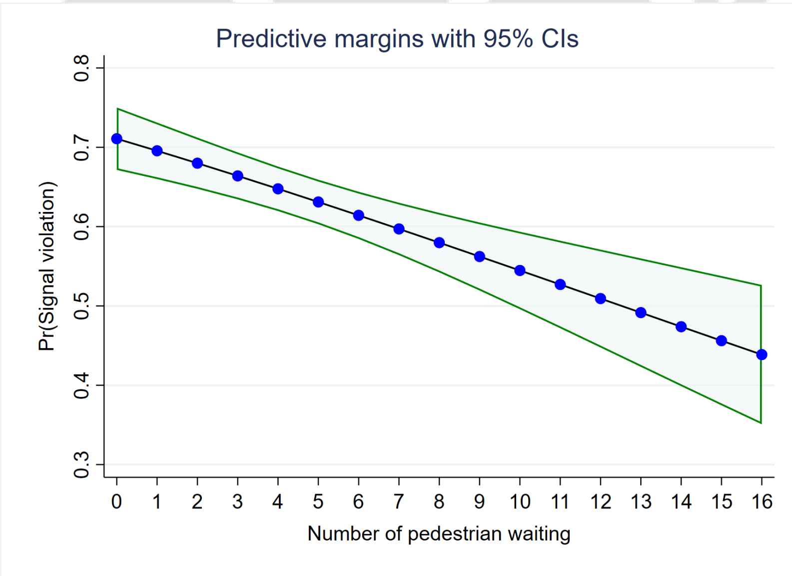


FIGURE 5.2: Predictive margins plot for number of people waiting upon arrival

In our investigation, as the group size (pedestrians already waiting or joining afterwards) increased, oncoming pedestrians avoided crossing the road in the red-light

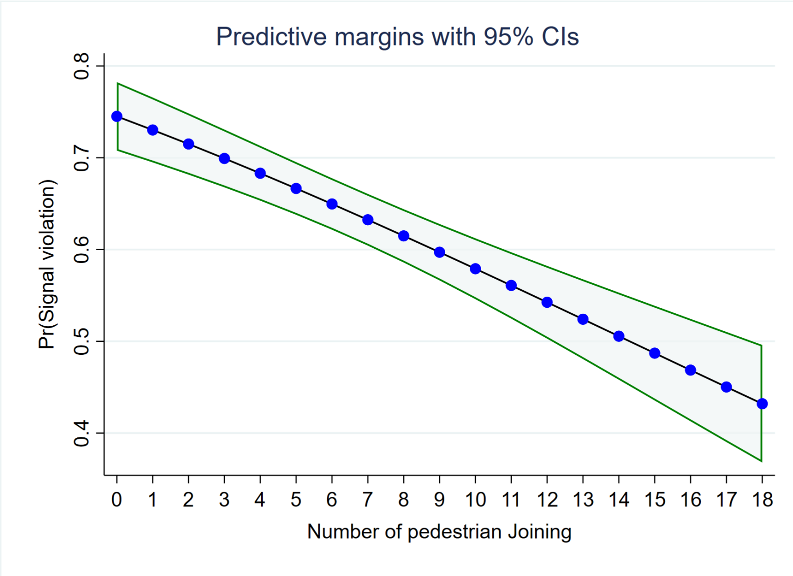


FIGURE 5.3: Predictive margins plot for number of people joining upon arrival

phase and decided to wait with the group, as they are likely to attribute their restraint to personal values rather than to gain/loss of considerations, implicating that social norms enhance the compliance with the law amongst pedestrians (Tyler T.R., 1990). Rosenbloom (2009) suggested that in a social framework, an individual's behavioural action can affect every other member of the group; thus, individuals deter themselves from taking unsocial steps that violate norms, rules, and regulations. Further, pedestrians reaching crosswalks in the red-light phase and standing alone are less concerned about social criticism (Pele et al., 2017; Rosenbloom, 2009); thus, their likelihood of crossing in the red-light phase is much higher. However, those who are surrounded by other pedestrians waiting for the green-signal phase feel more committed to social order, therefore sticking with the social norms (Wang et al., 2011). Although exceptions exist. Yagil (2000) studied beliefs, motives, and situational factors in road-crossing behaviour and concluded that other pedestrians' presence affects road-crossing behaviour, not because they serve as role models but because they stimulate conformity. Hyman et al. (2009) also explained that in a group setting, the "diffusion of responsibility" (Harrell, 1991) effect occurs when the subject delegates the task to other pedestrians and rely more on other group members as they believe that a partner, who is physically present, may help with the recognition of a significant stimulus.

5.8.4 The Presence of Signal Violators

"Hypothesis 2" divulgates that the likelihood of signal violation tendency would increase if the subject pedestrian observed a considerable number of pedestrians crossing from the same direction or the opposite direction during the red-light phase. The positive model coefficients revealed that this hypothesis is correct only when a significant number of pedestrians were observed crossing in the same direction, while pedestrians crossing from the other direction had no significant influence. The predictive margins plot (Figure 5.4) illustrates that the signal violation likelihood of a candidate pedestrian (who is waiting at the curb) increases when he/she observes that more and more neighbours crossing the road successfully violating the signal. The primary reason behind such behaviour could be that pedestrians crossing in the same direction provide a better representation of oncoming difficulties (risk of approaching traffic and available gap size) that pedestrians might face if they follow the red-light crossers (who were crossing in the red-light phase).

The significant positive coefficient also indicated that the count of pedestrians (observed in the current study) crossing in the red-signal phase is an appropriate predictor compared to the binary count (someone crossed in red/ did not cross) used by Rosenbloom (2009) in her study of single versus group crossing behaviour, where it was observed to be insignificant. When oncoming pedestrians observe more people completing the crossing successfully in the red-signal phase, it gives them a false sense of safety, making them believe in social information disseminated from red-light

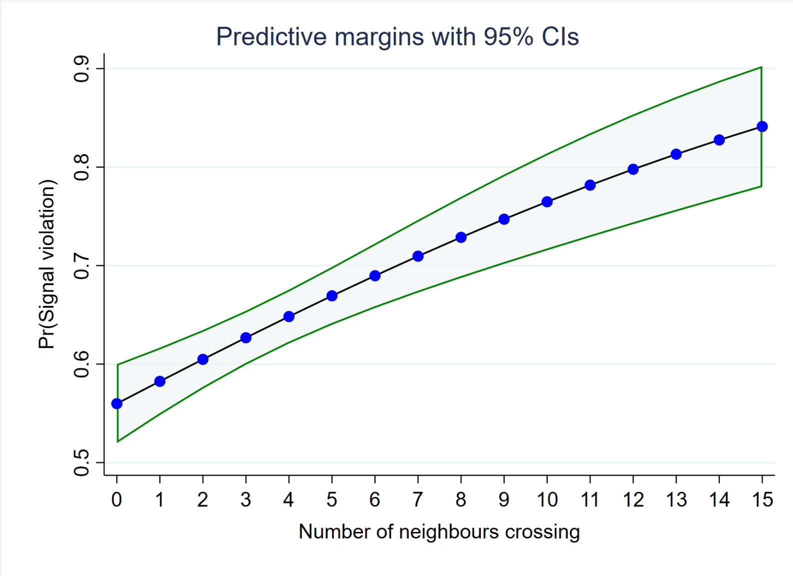


FIGURE 5.4: Predictive margins plot for number of neighbours crossing during red-light phase

crossers. Thus, people get encouraged to copy others' behaviour (start crossing in the red-signal phase), especially if they find that accepting risks provides satisfactory results in terms of a successful crossing. Bandura (1977, 1986) explained the motive behind such a risky decision, and reported that humans acquire the behaviour (red-light phase crossing) by observing others and the surrounding environment. Observing others' behaviour provides them with innovative ideas, and if the ideas are found to be satisfactory, it acts as a guide for future decision-making. Hence, risk-taking behaviour is not just initiated by self-motivated decisions but also from past accumulated experiences, i.e., by seeing others' successful/unsuccessful risky crossing choices.

Faria et al. (2010) studied the neighbours' effect and observed similar behaviour. They reported that pedestrians often try a few times to get familiar with the actual risk associated with crossing in the red-light phase before making the final crossing decision. For example, if a pedestrian found that the crossing risk was beyond their acceptable risk limits, they abort crossing and return to a safe position and update their decision with a better and safer crossing one. This is how pedestrians gain experience in different traffic scenarios, this experience works as a basis for future decision-making. Another explanation of this type of behaviour could be that people in a "group setting" feel safer. The concept of "safety in number" is also explained by Vujanić et al. (2014), who pointed out that pedestrians in a group feel safer even when crossing the road by taking the risk and believe that oncoming traffic will easily spot them and provide safe access for crossing.

5.8.5 The Waiting Time and Signal Cycle Length

Studies effectuated on waiting time reported a direct relationship between waiting time and signal violation. They reported that red-light phase length is the most crucial factor in models based on integrated field observations and questionnaire data (Yang and Sun, 2013). The longer the red-light phase length, the more likely to cross during the red-light phase. In situations where signal phase length becomes too long (due to heavy traffic flow and few adequate, acceptable gaps are available), the waiting time frame does not meet the pedestrians' expectations (exceeds threshold waiting time), and thus, they violate the signal.

The present violation model estimate confirmed the hypothesis that as the waiting time for safe crossing (green-phase crossing) gets longer, pedestrians lose patience and often violate signals. The 'waiting time till green initiation' variable estimate showed that in comparison to a waiting time of 0-20 seconds (reference category), when pedestrian encountered a longer waiting time of 21-40 seconds or >40 seconds, the signal violation probability increased by 18.3% and 28.4%, respectively. The current findings are similar to past studies conducted in Asian countries, which suggested keeping the red-light phase duration below 40 s. Investigation showed that pedestrians lost patience when red-light phase length exceeded 40 seconds and committed more signal violation (BAASS, 1989; Wang et al., 2011).

The impact of varying phase lengths can also be observed from the cycle length variable. The model estimate showed that in comparison to cycle length of 0 to 100 seconds, the pedestrians who arrived at the signal cycle category of 100-150 seconds (or >150 seconds) are 18.7% (or 40.4%) more likely to violate the signal.

Additionally, the model estimate showed that traffic volume significantly affected signal violation. As the traffic volume increases, pedestrians were less likely to take signal violation risk (Figure 5.5). The availability of safe gaps decreased with increase in traffic volume, discouraging pedestrians from accepting such unsafe gaps to violate the signal.

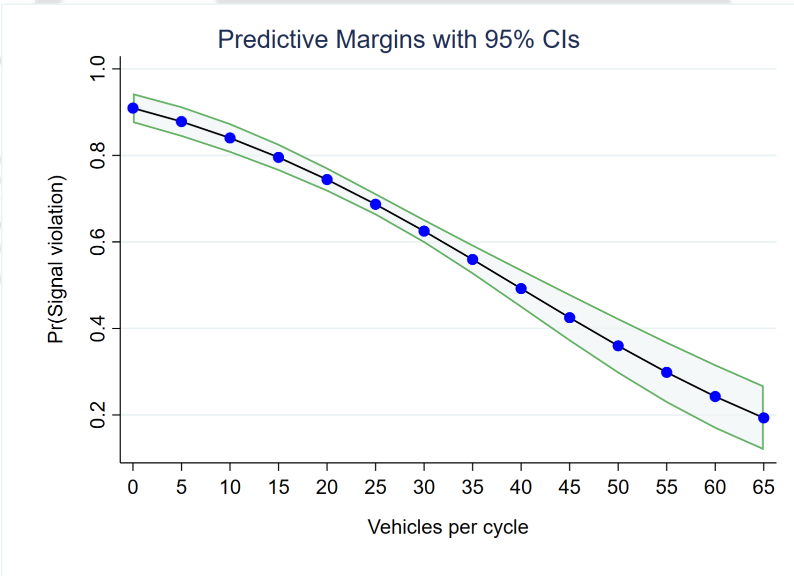


FIGURE 5.5: Predictive margins plot for number of vehicles per cycle

5.9 Practical Implications from the Results

In the current study, pedestrians' incidence of red-light phase violation behaviour was observed at three signalised intersection crosswalks (one-way) to identify various social and non-social factors that may affect red-light phase crossing behaviour. While illegal crossing might not be directly associated with substantial number of fatal crashes involving pedestrians, it is a significant indicator of riskier behaviour leading to injuries. Hence, understanding factors influencing such behaviour could be a crucial step towards a better road safety management. Based on the model findings, the following practical implications could be suggested:

- The study provided empirical evidence that pedestrian's red-light phase violation was influenced by social and non-social factors. This is a significant finding because both of those factors can be used to understand signal violation behaviour. The red-light violation could be utilised as a surrogate measure to identify unsafe intersections in Kolkata city.

- Regarding the group effect, the tendency to wait on the red-light phase was greater when more people were waiting at the curb, either when a pedestrian arrived or other pedestrians joined after the arrival of the subject. The present finding indicates a group's power to influence its members to obey the law positively. Media campaigns should utilise this power for educational purposes by emphasizing the positive value of social control and its benefits.

- An interesting finding emerged from this study: the chances of pedestrians' signal violation are significantly high if a pedestrian found that a significant number of neighbours successfully crossed the road, violating the red-signal phase.

- A unique finding from this study is that the number of neighbours crossing the street in the red-light phase has a significant influence on signal violation behaviour compared to the binary selection (someone crossed in the red-light phase or did not cross) used in past studies.

- The model outcome also suggested that if pedestrians experience a longer waiting time for safe crossing, they become impatient and engage in illegal crossing. The present finding points to the deficiencies in signal design and infrastructure planning. Hence, the grade-separated facility should be provided to segregate pedestrians and motorists completely in busy intersections where the traffic cycle length or red-light phase length is too long. In the absence of such facilities, strict enforcement could encourage pedestrians to comply with the traffic signal. Additionally, bringing legislation for restricting signal violations in the road environment into the future planning process through practical interventions could improve signal compliance.

5.10 Summary

The binary logistic regression model for estimating the probability of signal violation is presented in this chapter, which determined the effect of social and non-social information factors in mixed traffic conditions, which is one of the objectives stated in section 1.4 of Chapter 1. It is understood that signal violation at intersection crosswalks is highly influenced by non-social and social information. The model results also revealed a significant effect of waiting time till green, signal cycle length, traffic, and information disseminated from other pedestrians on signal violation likelihood. Some variables (age, group size, carrying luggage, crosswalk violation and number of pedestrians crossing from other direction) are eliminated due to their statistical insignificance at 90% or above. The impact of significant variables was explained using average marginal effects. The next chapter discusses the impact of pedestrian distraction on road crossing behaviour at signalised intersection crosswalks.

Chapter 6

Understanding Distracted Road Crossing Behaviour at Signalized Intersections

6.1 General

Pedestrian-distracted road crossing behaviour has been a major concern for road safety researchers. Without accident statistics, validating distraction-related injuries and deaths in road environments is impossible in developing countries like India. Thus, to understand the unsafe behaviour displayed by distracted pedestrians at crosswalks, initially, a questionnaire-based survey was conducted across eight intersections in Kolkata (India). The results revealed that 13.7% (61) of the respondents were subjected to at least one near-miss, and 4.5% (20) experienced a crash in the past (assuming the pedestrian encountered just an injury). However, the total number of pedestrian deaths due to Road Traffic Injuries in India is around 35%

(52996) of all road traffic deaths, 1,51,417 (death as per 2018 official estimates). Assuming a 4.5% crash rate due to pedestrian distraction (taking the conservative side, i.e., considering just crashes involving injuries), it comes out to around 2285 (52996×0.045) crashes, which is still a significant number. Therefore, to reduce the number of crashes involving distracted pedestrians, understanding distracted road crossing behaviour is essential in developing countries like India. The current study tried to fill this gap.

6.2 Analysis Methodology

The primary objective of this study is to identify the factors associated with distracted road-crossing behaviour. Pedestrian distraction is binary by nature as it contains two possible outcomes, such as mobile phone distraction versus no distraction or texting distraction versus no distraction and so on. As such, binary logistic regression models were fitted to analyse these types of crossing behaviour, as this approach models the event probability for a categorical response variable with two outcomes. In the case of the binary logistic regression model, the probability of distraction was estimated using the following equation (Eq. 6.1):

$$P_i = \frac{\exp^{(\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i}...+\beta_k X_{k,i})}}{1 + \exp^{(\beta_0+\beta_1 X_{1,i}+\beta_2 X_{2,i}...+\beta_k X_{k,i})}} \tag{6.1}$$

where P_i is the probability of pedestrian i crossing distracted. β_0 is the model constant, and $\beta_1,\ldots,\beta_k$ are coefficients estimated by maximum-likelihood from the corresponding explanatory variables $X_1 \ldots X_k$. The logistic regression was estimated using the statistical programming language R (version 4.1.1). Average Marginal Effects (AME) were calculated at every observed value of observation and averaged across the results (AMEs) using R's "margin" library (Leeper, 2018b) and reported in respective tables.

6.3 Association Among Variables

Identifying multicollinearity among variables using a pairwise-correlation estimation process (before model fitting) is a common practice to eliminate predictors having a correlation value beyond a threshold limit (rule of thumb, r >0.4). In the present study, as most of the variables are of categorical type, thus instead of estimating pair-wise Pearson correlation coefficient (r), Theil's U pair-wise association statistics was estimated (Theil, 1958, 1966). Theil's U statistics gives an association strength for categorical variable pairs ranging from 0 to 1, where 0 indicates no association and 1 indicates perfect association. The Theil's U is an asymmetric measure, [U (x,y) $\neq$ U(y,x), where U is Theil's U]. The pairwise Theil's association estimates are reported in Figure 6.1. The estimated pair-wise association strength among categorical variables was found to be weak (Theil's U: $\leq$ 0.27), indicating that the selected variables are weakly associated with each other and can be used for the model estimation.

6.4 Results and Discussion

One-way crosswalks data is utilised (n = 2360) to understand the pedestrians' distracted road crossing behaviour. Six separate binary logistic regression models (mobile talking, texting, headphones use, eating/drinking/smoking, group talking, and

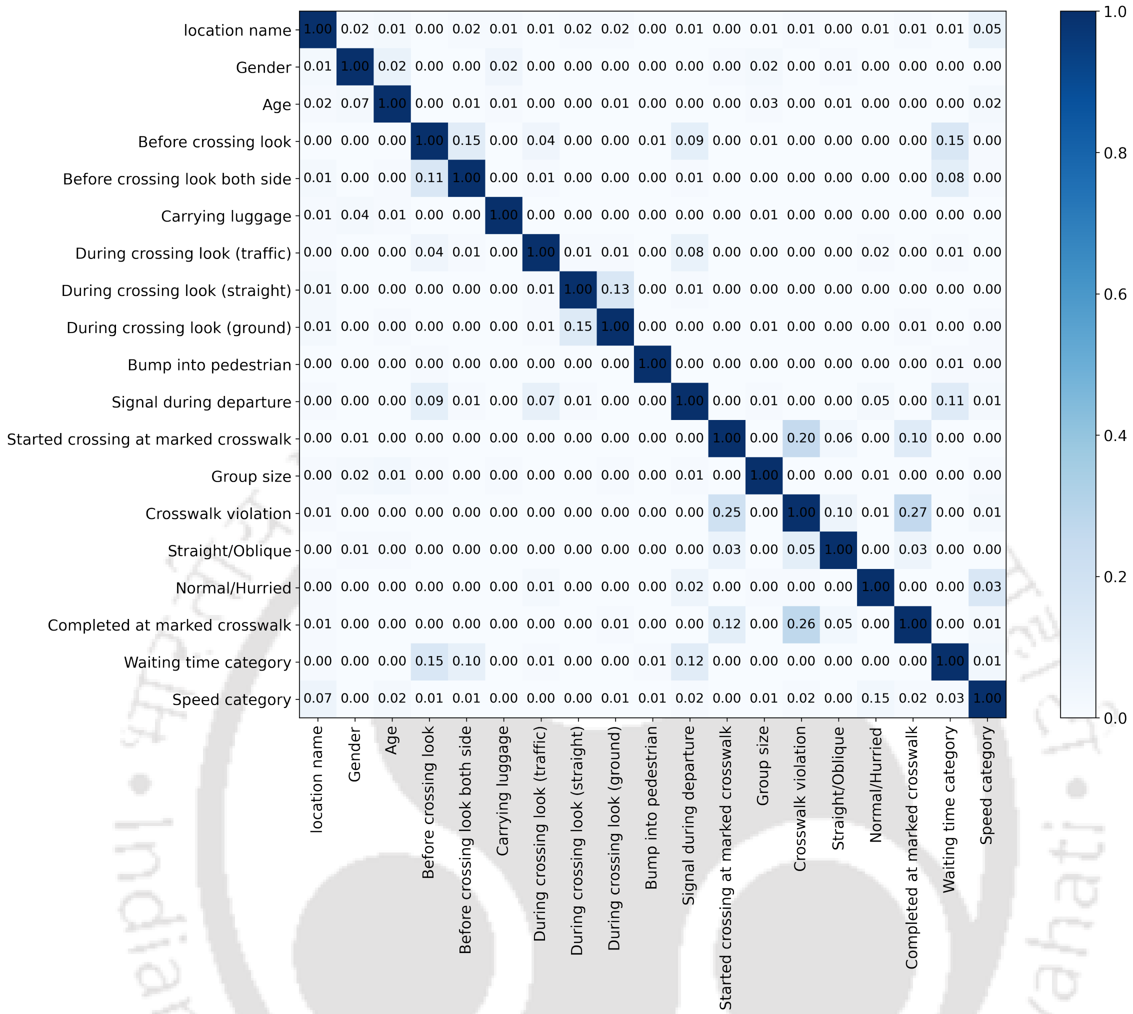


FIGURE 6.1: Theil's U pair-wise association plot.

holding a phone in hand) were estimated. As discussed in Chapter 4, site, demographic, behavioural, violation, state of crossing, and cautionary behaviour variables were included in the modelling process. A step-wise elimination process was adopted for variable selection while fitting all binary logistic regression models. In the step-wise elimination process, the model was initially fitted with all variables and subsequently, certain variables of the models were removed based on the judgement whether the removal of that caused a significant change in other variables or the aggregate model. The iterative process goes on until there is no further improvement. The threshold value for confidence level was set to 90% or higher (p-value

$\leq$ 0.10). The subsequent paragraphs describe model results in detail and highlight the different types of distraction influencing road crossing behaviour at intersection crosswalks. Table 6.1-6.6 presents the fitted binary logistic regression model summaries to identify the factors influencing the distracted road crossing behaviour.

6.4.1 Talking on Mobile Phone

The results of the mobile talking-related distraction model summary (Table 6.1) revealed that pedestrian talks on mobile more at GPO location (Kolkata administration zone) and relatively less at "BBG" when compared to "Dalhousie Square". It is also found that young adult pedestrians (18 to 29 years) talk on the phone more compared to the old pedestrian (>60 years) age group. Further, pedestrians distracted by mobile phone talking performed unsafe behaviour. Results highlighted that pedestrians are less likely to (i) look at traffic before starting to cross, (ii) observe their surroundings (both sides for extra caution) for potential traffic threats and (iii) start crossing on the marked crosswalk and on the contrary more likely to complete crossing on the marked crosswalk, these were expected results similar to prior research in this area (Pešić et al., 2016). Single pedestrians were found to use their mobile more frequently for talking compared to those in a group of two or more, which might be due to the lack of social interaction of a single pedestrian and being subjected to boredom. It is also observed that during road crossing, the pedestrians talking over mobile were found to perform less cautionary behaviour, i.e., they were less likely to watch for oncoming traffic and even less frequently move their head to be aware of oncoming traffic information. Mobile phone talkers were about 4.5% more likely to bump into oncoming pedestrians, which was expected and consistent with past research (Horberry et al., 2019). The reason might be that mobile phone talking increases cognitive workload, resulting in diminished situational awareness.

Finally, compared to pedestrians waiting for at least 1 second after arrival, those who immediately start crossing were 6% less likely to talk on their mobile phone.

Table 6.1: Mobile phone taking distraction model estimates

Features	Variable name	Coefficient (β)	Standard error	P-value	AME
Site	Constant	-0.370	0.392	0.345	na
	Site 2 (BBG)	-0.119	0.198	0.549	-0.009
	Site 3 (GPO)**	1.019	0.201	<0.001	0.114
Demographic	Age <18	-0.644	1.097	0.557	-0.030
	Age 18-29**	1.420	0.334	<0.001	0.139
	Age 30-45**	0.746	0.283	0.008	0.058
	Age 46-60**	0.611	0.279	0.028	0.046
Pedestrian behavioural characteristics	Group: 2 or more pedestrians**	-1.272	0.487	0.009	-0.082
	Waiting time: 0**	-0.601	0.199	0.003	-0.060
Violation	Started crossing at marked crosswalk: yes**	-0.513	0.185	0.006	-0.052
	Completed crossing on marked crosswalk: yes**	0.365	0.178	0.039	0.033
State of crossing	Carrying Luggage: yes**	-0.466	0.161	0.004	-0.044
	Crossing path: oblique**	0.554	0.240	0.021	0.059
Cautionary Behaviour and Critical Events	Looked at traffic before crossing: yes**	-0.785	0.189	<0.001	-0.079
	Looking Both side before crossing: yes**	-0.617	0.249	0.013	-0.051
	Looked straight during crossing: yes**	-1.011	0.182	<0.001	-0.110
	Looked at traffic during crossing: yes*	-0.517	0.284	0.068	-0.049
	Frequency of looking at traffic while crossing*	-0.330	0.169	0.051	-0.031
	Bumped into pedestrian: yes*	0.432	0.256	0.091	0.045

Model Fit Statistics
McFadden (Pseudo R-squared) = 0.132

Likelihood ratio test
Df.diff = -18, LogLik.diff = -83.106, χ^2 = 166.21 and P-value = <0.001

Reference levels
Site: Site1-Dalhousie; Age: >60; Group size: Single; Waiting: one or more (s); Started or completed crossing at marking: no; Carrying luggage: no; Crossing path: perpendicular; Looked at traffic before crossing: no; Looked on both side: no; Looked straight/traffic: no; Bumped: no.

Note:

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

6.4.2 Texting on the Phone

Observations on "texting on the phone" results (Table 6.2) highlighted that the young age group (<18 years) was more likely to text than any other age group. Pedestrians who text during crossing were less likely to, (i) look at traffic before starting to cross, (ii) look straight during the crossing, and (iii) obeying crosswalk marking. Pedestrians crossing alone were found to be more engaged in texting than in-group, which might yet again be due to boredom arising mainly from lack of social interactions. Finally, the crossing speed variable highlighted that pedestrian engaged in texting maintain overall slow (<1.2 m/s) crossing speed compared to pedestrians with no distraction, which might cause late finishing or finishing on the red pedestrian signal phase that might increase traffic exposure risk.

TABLE 6.2: Texting distraction model estimates

Features	**Variable name**	**Coefficient (β)**	**Standard error**	**P-value**	**AME**
	Constant**	-5.923	0.947	<0.001	na
Demographic	Age <18**	3.764	0.944	<0.001	0.162
	Age 18-29**	3.625	0.771	<0.001	0.147
	Age 30-45**	2.748	0.741	<0.001	0.073
	Age 46-60**	1.766	0.753	0.019	0.028
Pedestrian behavioural characteristics	Signal during departure: Red**	1.354	0.666	0.042	0.079
	Signal during departure: Flash	1.262	1.408	0.37	0.072
	Group: 2 or more pedestrians**	-1.529	0.693	0.027	-0.044
	Crossing speed: ≤1 (m/s)**	0.996	0.42	0.018	0.045
	Crossing speed: >1-1.2 (m/s)**	0.969	0.364	0.008	0.043
	Crossing speed: >1.2-1.4 (m/s)	0.46	0.369	0.213	0.017
	Crossing speed: >1.4-1.5 (m/s)	0.16	0.462	0.728	0.005
Violation	Started crossing at marked crosswalk: yes	0.435	0.339	0.200	0.018
	Completed crossing on marked crosswalk: yes	0.567	0.352	0.107	0.024
	Crosswalk violation: no**	-0.929	0.389	0.017	-0.044
Cautionary Behaviour	Looked at traffic before crossing: yes**	-1.123	0.266	<0.001	-0.056
	Looked straight during crossing: yes**	-0.618	0.269	0.022	-0.03
	Looked at ground during crossing: yes**	1.864	0.379	<0.001	0.065
	Frequency of looking at traffic while crossing**	-0.551	0.171	0.001	-0.025

Model Fit Statistics
McFadden (Pseudo R-squared) = 0.233

Likelihood ratio test
Df.diff = -18, LogLik.diff = -80.99, χ^2 = 161.98 and P-value = <0.001

Reference levels
Age: >60; Departure signal: green; Group size: Single; Crossing speed: >1.5; Started or completed crossing at marking: no; Crosswalk violation: yes; Looked at traffic before crossing: no; Looked straight/ground during crossing: no.

Note:

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

However, a study conducted in New York City and Arizona State did not find any significant association between texting and walking speed (Russo et al., 2018); while other studies highlighted that texting might cause crossing delay and operational consequences (Chen and Pai, 2018; Gillette et al., 2016; Mohd Syazwan et al., 2017).

6.4.3 Headphones Use

The Table 6.3 model estimates drawn from the headphones use revealed that females were more likely to use headphones compared to males when crossing the road. Pedestrians using headphones exhibited less cautionary behaviour; headphone users were less likely to look straight ahead and even less likely to turn their heads for traffic updates.

The reason behind such behaviour could be use of headphones captures attention of pedestrians while obstructing traffic noise. This reduces the existing traffic awareness. Further, it makes pedestrians to envision being in a natural environment, withdrawing their alertness from the situation (Walker et al., 2012). Finally, headphone users were more likely to follow an oblique path (instead of the shortest/straight path) within a zebra crossing.

6.4.4 Eating/Drinking/Smoking

The results of eating, drinking, and smoking model estimates (Table 6.4) revealed many insignificant variables. The model estimates revealed that female pedestrians were less likely to eat/drink/smoke compared to male pedestrians while crossing at an intersection crosswalk. In addition, pedestrians with such distraction perform less cautionary behaviour, i.e., less frequently be attentive to traffic while crossing

TABLE 6.3: Headphones distraction model estimates

Features	**Variable name**	**Coefficient (β)**	**Standard error**	**P-value**	**AME**
	Constant**	-4.115	0.797	<0.001	na
Demographic	Gender: female**	1.525	0.464	0.001	0.039
Pedestrian behavioural characteristics	Group: 2 or more pedestrians	-0.918	1.044	0.379	-0.010
Violation	Crosswalk violation: no**	1.469	0.517	0.004	0.021
State of crossing	Crossing path: oblique**	1.460	0.513	0.004	0.038
Cautionary Behaviour	Looked straight during crossing: yes*	-0.925	0.530	0.081	-0.018
	Looked at ground during crossing: yes	-0.675	0.490	0.168	-0.011
	Looked at traffic during crossing: yes*	-0.760	0.423	0.072	-0.013

Model Fit Statistics
McFadden (Pseudo R-squared) = 0.112

Likelihood ratio test
Df.diff = -7, LogLik.diff = -14.351, χ^2 = 28.703 and P-value = <0.001

Reference levels
Gender: male; Group size: Single; Crosswalk violation: yes; crossing path: perpendicular; Looked at traffic/ground/straight during crossing: no.

Note:

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

the road and even present slow crossing speed, which might result in late finishing or missing the best time for crossing, an expected result also consistent with past research (Zhou et al., 2019).

6.4.5 Group Talking

In addition to digital distraction, social distractions such as crossing the road in a group and simultaneously talking with group members might result in unsafe road crossing. Table 6.5 clearly highlights that significant group talking behaviour was observed in GPO compared to Dalhousie square intersection. According to the model estimates, pedestrians conversing in groups were less likely to (i) look straight

TABLE 6.4: Eating/Drinking/Smoking distraction model estimates

Features	**Variable name**	**Coefficient (β)**	**Standard error**	**P-value**	**AME**
	Constant **	-5.336	0.789	<0.001	na
Site	Site 2 (BBG)	-0.355	0.331	0.283	-0.012
	Site 3 (GPO)	0.418	0.323	0.194	0.020
Demographic	Gender: female*	-1.399	0.735	0.057	-0.035
Pedestrian behavioural characteristics	Waiting time: 0**	1.184	0.387	0.002	0.037
	Crossing speed: $\leq$1 (m/s)**	2.382	0.773	0.002	0.06
	Crossing speed: >1-1.2 (m/s)**	2.300	0.750	0.002	0.055
	Crossing speed: >1.2-1.4 (m/s)**	2.157	0.744	0.004	0.048
	Crossing speed: >1.4-1.5 (m/s)	0.503	1.008	0.618	0.004
State of Crossing	Carrying Luggage: yes	-0.412	0.253	0.104	-0.017
Cautionary Behaviour	Frequency of looking at traffic while crossing**	-0.384	0.155	0.013	-0.016

Model Fit Statistics
McFadden (Pseudo R-squared) = 0.115

Likelihood ratio test
Df.diff = -10, LogLik.diff = -33.551, χ^2 = 67.102 and P-value = <0.001

Reference levels

Site: Site 1-Dalhousie sq.; Gender: male; Waiting time: one or more (s); Crossing speed:>1.5; Carrying luggage: no.

Note:

****** Denotes variable significance at 95% confidence level;

***** Denotes variable significance at 90% confidence level.

ahead and watch for oncoming opposite directional pedestrians, (ii) look towards the ground, and less frequently observe traffic before attempting to cross, which might be due to the fact that while engaged in talking, two pedestrians might frequently look towards each other than the surrounding. Further, the majority of the group talking pedestrians were crossed the road, maintaining 1 to 1.2 m/s crossing speed.

6.4.6 Holding Phone in Hand

'Holding the phone in hand' might not be life-threatening. However, it might cause an increased temptation of looking into the phone when a notification was received.

TABLE 6.5: Group talking distraction model estimates

Features	Variable name	Coefficient (β)	Standard error	P-value	AME
	Constant**	-1.149	0.315	<0.001	na
Site	Site 2 (BBG)	0.094	0.194	0.626	0.008
	Site 3 (GPO)**	0.871	0.209	<0.001	0.100
Pedestrian behavioural characteristics	Crossing speed: ≤1 (m/s)	0.437	0.305	0.152	0.036
	Crossing speed: >1-1.2 (m/s)**	0.934	0.268	<0.001	0.091
	Crossing speed: >1.2-1.4 (m/s)**	0.650	0.267	0.015	0.057
	Crossing speed: >1.4-1.5 (m/s)*	0.599	0.325	0.065	0.052
Cautionary Behaviour	Before crossing look frequency**	-0.282	0.075	<0.001	-0.029
	Looked straight during crossing: yes**	-1.226	0.185	<0.001	-0.155
	Looked at ground during crossing: yes**	-0.632	0.172	<0.001	-0.067

Model Fit Statistics
McFadden (Pseudo R-squared) = 0.101

Likelihood ratio test
Df.diff = -9, LogLik.diff = -69.119, χ^2 = 138.24 and P-value = <0.001

Reference levels

Site: Site1-Dalhousie sq.; Crossing speed: >1.5; Looked at ground/straight during crossing: no.

Note:

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

Similar to a previous investigation (Baswail et al., 2019), in the current study, holding a phone was considered and treated as one form of distraction. However, many variables in the analysis did not make good significance (refer Table 6.6), indicating overall less difference in pedestrian behaviour of holding a phone versus no phone. Results highlighted, compared to Dalhousie square intersection, BBG had a higher number of pedestrians holding a phone while crossing. Female pedestrians were more likely to hold a phone compared to male pedestrians, which is mainly happened due to the Indian female attire. Young pedestrians were more likely to hold phones as studies showed that young pedestrians check their devices more frequently for updates even when there was no prompt for any incoming message or call (Gold et al., 2015).

Table 6.6: Holding phone in hand distraction model estimates

Features	Variable name	Coefficient (β)	Standard error	P-value	AME
	Constant**	-3.859	0.415	<0.001	na
Site	Site 2 (BBG)**	0.693	0.201	<0.001	0.057
	Site 3 (GPO)	0.185	0.248	0.456	0.013
Demographic	Gender: female**	1.227	0.214	<0.001	0.135
	Age <18**	1.709	0.689	0.013	0.113
	Age 18-29**	2.052	0.439	<0.001	0.158
	Age 30-45**	1.623	0.406	<0.001	0.103
	Age 46-60	0.586	0.425	0.168	0.023
Pedestrian behavioural characteristics	Group: 2 or more pedestrians	-0.56	0.367	0.127	-0.038
State of crossing	Crossing pace: hurried**	-0.804	0.405	0.047	-0.05
Cautionary behaviour	Looked both side before crossing: yes	-0.326	0.234	0.164	-0.024

Model Fit Statistics
McFadden (Pseudo R-squared) = 0.102

Likelihood ratio test
Df.diff = -10, LogLik.diff = -54.229, χ^2 = 108.46 and P-value = <0.001

Reference levels
Site: Site1-Dalhousie sq.; Gender: male; Age: >60; Group size: Single; Crossing pace: normal; Looked both side before crossing: no.

Note:
** Denotes variable significance at 95% confidence level;
* Denotes variable significance at 90% confidence level.

6.5 Practical Implications from the Results

- Overall, the study results suggested that the young age (<29 years) pedestrians are prone to more digital distraction. Hence, Policies need to be more focused on digital distraction that involves young pedestrian age groups. Interventions to reduce digital distraction need to be implemented near the school zones and areas with more young residents.

- Survey response showed that among all reported mobile use, the primary segment of the respondents was reported to be distracted from frequent mobile phone talking (81.2%) and when it was work-related (41%).

- Additional interventions to attract the pedestrian attention for safe crossing need to be deployed. Innovative interventions, such as embedding LED lights/ audio-based systems in pathways could be used to seek pedestrians' attention, and warn them about signal status (Larue et al., 2019), which would indirectly increase the cautionary (head movement) behaviour and eventually reduce distraction during road crossing.

- Machine Learning (ML) and computer vision-based interventions could be deployed for pedestrian safety improvements, such as integrating warning mechanisms in mobile applications that sense potential distraction during road crossing and warn pedestrians about oncoming traffic threats could be used to improve crossing safety.

- The distraction themed questionnaire survey results presented in Section 4.9.1 (Chapter 4) revealed that about 13.7% (61) of survey respondents were subjected to at least one near-miss in the past, and 4.5% (20) experienced at least one accident. This indicates that pedestrians' safety do get compromised when they cross distracted. Thus, intersections with high share of distracted pedestrians require especial attention. For example, flashing LED lights and Variable Message Sign (VMS) could be installed to reduce serious vehicle-pedestrian conflicts and improve drivers' yielding behaviour at intersection crosswalks (Hussain et al., 2021).

- Publicity campaigns, for example, advertising on vehicles, campaigns of road crossing safety, and posters with warnings citing the danger involved in distracted road crossing could help in minimising such behaviour.

- Bringing legislation for restricting phone use in the road environment into the future planning process in the form of practical intervention could reduce the pedestrian's distracted road crossing behaviour (Osborne et al., 2020). For example, fines/community service could be imposed on pedestrians if, it were deemed, they were putting themselves or others at risk from their phone use.

6.6 Summary

The models presented in this chapter are aimed at identifying significant factors that influence distracted road crossing behaviour. The gender, age, glance, carrying luggage, crossing speed and violation found to be significant. The results revealed that pedestrians of young age group frequently engaged in digital distraction. Pedestrians talking on a phone were more likely to 'nearly hit/bump' into another oncoming pedestrian. In addition, distracted pedestrian less likely to perform 'glance/cautionary behaviour' before as well as during crossing. Further, pedestrians who texts were more likely to cross in red-signal phase than pedestrians without distraction. The next chapter discusses pedestrian waiting behaviour at signalised intersections using time to event analysis (survival analysis).

Chapter 7

Waiting Behaviour Modelling using Hazard-Based Duration Approach

7.1 General

The pedestrian waiting duration at the intersection crosswalk was found to be an important factor in road-crossing decision-making. It is often found related to the allotted pedestrian red-phase length at signalised intersections. Thus, an understanding of pedestrian waiting behaviour is required to identify the optimal red-phase length that enhances pedestrian signal compliance. In the present chapter, the waiting duration of pedestrians crossing through the intersection crosswalks was modelled using a hazard-based duration approach. Primarily, a non-parametric hazard model was utilised to model the survival probabilities of duration data using a Kaplan-Meier survival curve. Further, a semi-parametric Cox-proportional hazard model was utilised to establish the relationship between the covariates and duration data. The proportional assumption was tested using a goodness-of-fit test. Following the non-conformity of proportional assumption, AFT methods (Weibull, Exponential, Log-normal, and Log-logistic) were fitted and compared to identify the best-fitted model that shows the designated risk related to each parameter.

7.2 Models and Algorithms

The models and algorithms include a brief description of different time-to-event modelling approaches, which include the non-parametric Kaplan-Maier (K-M) estimator, semi-parametric Cox Proportional Hazard (Cox-PH) regression and Accelerated Failure Time (AFT) model.

7.2.1 Survival Analysis

Survival time, also known as failure time, is the amount of time it takes from the start of an operation to the end. Complete (uncensored) data and censored data are two types of survival time data. The time when the detected object appeared (end event) and the time when the information collected is complete is referred to as complete data. If the follow-up is stopped for some reason, the answer (end event) has not yet been observed, and the time information reported is incomplete.

Pedestrian waiting times were divided into censored and uncensored data in this analysis. The uncensored data describes the occurrence of the event, here, for example, when a pedestrian stops waiting to cross the intersection during the red light signal (violate signal). Otherwise, it is called censored data if a pedestrian waits before the green light appears. The pedestrian survival function is the proportion of pedestrians who remain waiting after time t. The survival function represents the likelihood that the person survives to time t (experiencing an incident after time t) as described in Eq. 7.1.

$$S(t) = P(T > t) = \int_t^{\infty} f(u)\, du \tag{7.1}$$

where $S(t)$ is the pedestrian's survival function, $f(u)$ is the probability density function, and T is the pedestrian's waiting time. There are n time data (including censored and uncensored). The data are organised from small to large as $t(1) \leq t(2) \leq \ldots \leq t(n)$, assuming that the events of the measurement object are independent of each other in each unit cycle and the survival probability is $p_1, p_2, p_3.....p_n$ then the survival function can be retrieved using the Eq. 7.2.

$$S(t) = p_1.p_2.p_3 \ldots p_t = \prod_{t_{i \leq t}} p_i \tag{7.2}$$

As a result, the cumulative survival probability, also known as the survival function, is about the time t survived and is regarded as the cumulative result of the previous t duration.

7.2.2 Non-Parametric Hazard Model

The KM analysis procedure, introduced by Kaplan and Meier in 1958, is used to estimate survival functions in survival analysis (Kaplan and Meier, 1958). This is a non-parametric modelling approach that involves the product limit of conditional probabilities for estimating survival function. The general equation for the Kaplan-Meier survival probability curve where $T \geq 0$ is a random variable is shown below:

$$S(t) = \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \tag{7.3}$$

t_i = time when at least one signal violation (event) occurs.

d_i = the number of signal violations (events) that took place in time t_i.

n_i = the pedestrians who waited (survived) up to time t_i.

The Kaplan-Meier curve consists of a step function which starts with 1 (at the beginning, the survival probability is 100%) and steps down with decreasing survival probability to zero.

7.2.3 Cox Proportional Hazard Model

The Cox hazard-based model, a semi-parametric model, is used to investigate the various factors that influence the survival time (Cox, 1972). The Cox hazard-based model has general regression analysis characteristics. It may compare different levels of a factor on the effect of survival time in the case of given other factors. It works in the same way as the logistic regression model. After calculating the regression coefficient, the relative risk of the corresponding factor can be calculated. However, the logistic regression model's dependent variable may only be a qualitative variable, such as the outcome of a case (end event), without accounting for the survival time. The Cox hazard-based model's dependent variables are survival time and the conclusion of events (Eq. 7.4). Here, t represents the time (waiting time) it takes for a pedestrian to cross a signalised intersection, and X represents several contributing factors. The Cox hazard-based model's basic expression type is:

$$h(t, X) = h_0(t)e^{(\beta_1 X_1 + \beta_2 X_2 + \ldots \beta_p X_p)} = h_0(t)e^{(\beta x)} \tag{7.4}$$

where $h(t, X)$ is the risk rate at the time in the case of given covariates. The reference risk function, $h_0(t)$, has a value of 0 for each vector. The regression coefficients referring to the determining variables are $\beta_1, \beta_2, \beta_3, \ldots, \beta_p$.

If $\beta_i > 0$, the risk rises as X_1 rises, indicating that X_1 is positively associated with the risk rate and that X_1 is a risk factor. If $\beta_i < 0$, the risk decreases as X_1 increases, indicating that X_1 is negatively related to the risk rate and that X_1 is a protective factor. If $\beta_i = 0$, the danger rate stays unchanged as X_1 increases, implying that X_1 has no impact on the pedestrian's waiting time. In addition, hazard ratios can be estimated, which provides a meaningful interpretation of the covariates. A hazard ratio is defined as the ratio of the hazard calculated for one individual characteristic and the hazard calculated for another individual characteristic (e.g., the hazard ratio for gender = hazard for male/hazard for female). The value of the hazard ratio ranges from zero to infinity and is calculated as:

$$Hazard\ Ratio = e^{coefficient\ value} \tag{7.5}$$

The hazard ratio (HR) is interpreted as:

HR = 1: No effect

HR < 1: Reduction in the hazard

HR > 1: Increase in hazard

For example, if the hazard for male pedestrians is thrice of female pedestrians, it indicates that the rate of signal violation is three times higher in male pedestrians than female pedestrians.

7.2.4 Parametric AFT Model

A parametric hazard model assumes that the survival event follows a known distribution. AFT models could be used when proportional hazard assumption criteria are not fulfilled. The AFT model assumes that covariates have a multiplicative effect on the survival time. The accelerator factor ($e^{coefficient}$) compares the decrease or increase in survival time corresponding with any constraint value of $s(t)$.

Suppose we have two populations, P and Q, with different survival functions, $S_P(t)$ and $S_Q(t)$, and they are related by some accelerated failure rate, λ:

$$S_P(t) = S_Q\left(\frac{t}{\lambda}\right) \tag{7.6}$$

For example, if λ for male pedestrians versus female pedestrians is 3, it is said that the survival of female pedestrians is thrice of male pedestrians. Additionally, this model has other distinct properties: the average survival time of population B is λ times that of population A. Similarly, with the median survival time. More generally, the λ can be modelled as a function of available covariates:

$$S_P(t) = S_Q\left(\frac{t}{\lambda(x)}\right) \tag{7.7}$$

$$\lambda(x) = exp\left(b_0 + \sum_{i=1}^{n} b_i x_i\right) \tag{7.8}$$

where this model can accelerate or decelerate failure times (in the current study, it indicates waiting times till signal violation) depending on subjects' covariates. The most commonly used distributions for evaluating the significant factors are Weibull, Exponential, Log-Normal, and Log-Logistic. The explanation of the Log-normal model is presented in the following paragraphs.

Log-Normal AFT model: If ϵ_i follows the standard normal distribution, then T_i is log-normally distributed. The density function of the normal distribution is:

$$f_{\epsilon i} = \frac{1}{\sqrt{2\pi}} exp - \left(\frac{\log t - \mu - \beta_1 x_1 - \ldots \beta_p x_p^2}{\sigma} \right) /2 \tag{7.9}$$

where: $x_1, \ldots, x_p$ are the explanatory variables with the coefficients $\beta_1, \ldots, \beta_p$; ϵ_i is the residual or unexplained variance in the log-transformed survival times; μ is the intercept and σ is the scale parameter respectively.

The survival function of the normal distribution is:

$$S_{\epsilon i} = 1 - \phi(\epsilon) \tag{7.10}$$

The distribution function for normal distribution is:

$$\phi(\epsilon) = \frac{\log t - \mu - \beta_1 x_1 - \ldots \beta_p x_p}{\sigma} \tag{7.11}$$

The cumulative hazard function can be written as:

$$H_{\epsilon i}(\epsilon) = -\log 1 - \phi(\epsilon) \tag{7.12}$$

Further, the hazard function is:

$$h_{\epsilon i}(\epsilon) = \frac{f_{\epsilon i}(\epsilon)}{S_{\epsilon i}(\epsilon)} \tag{7.13}$$

7.2.5 Model Goodness of Fit Statistics

The existing study uses three evaluation criteria to identify the best-suited AFT model. The evaluation criteria are the likelihood ratio test, Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). The values for each model were compared to select the best-fitted model.

The Likelihood Ratio Test: Likelihood ratio tests are used to compare the goodness of fit of two statistical models. The LRT compares two hierarchically nested models to determine whether or not adding more parameters (adding complexity) makes the model significantly more accurate. The likelihood ratio is estimated by taking the ratio of likelihoods of two models. It is computed using the following equation:

$$X_L = -2\left(LL(\beta_0) + LL(\hat{\beta})\right) \tag{7.14}$$

Where $LL(\beta_0)$ is the likelihood of a null model when all parameters are set to zero and $LL(\hat{\beta})$ is the log-likelihood of the model with all fitted parameters on convergence. The log-likelihood ratio test statistics and p-value of statistical significance of 0.05 ensure that at least one significant parameter exists in the model.

Akaike Information Criteria (AIC): AIC metric is used to compare semi-parametric and parametric models. The AIC is proposed by Akaike (1974). It is a measure of the goodness of fit of a statistical model. The AIC is calculated as:

$$AIC = -2 \ln L + 2(P + K) \tag{7.15}$$

Where $\ln L$ is the maximised log-likelihood of the model, P is the number of parameters estimated, and K is the number of coefficients (excluding constant) in the model. The model with the smallest AIC value is considered the best-fitted model.

Bayesian Information Criteria (BIC): BIC is a criterion functioning similar to AIC (Schwarz, 1978). The BIC is calculated as:

$$BIC = -2\ln L + (P + K) \times \ln N \tag{7.16}$$

where $\ln L$ is the maximised log-likelihood of the model, P is the number of parameters estimated, K is the number of coefficients, and N is the sample size. A lower BIC value represents a better model fit.

7.3 Results

7.3.1 Signal Violation Model Results

In total, 2800 pedestrians were observed at eight selected crosswalks. The observation consists of 1930 males and 870 females. The age group composed of 196 young (<18: 7%), 720 young adults (18-29: 25.7%), 1049 adults (30-45: 37.5%), 681 old adults (46-60: 24.3%) and 154 elderly (>60: 5.5%) pedestrians. Out of the total pedestrians, 2089 (74.6%) arrived at the intersections during the red-signal phase. These red-signal phase arriving pedestrians were separated and considered for the final analysis. One thousand five hundred thirty-four pedestrians violated the signal (73.43%) who arrived at the red-signal phase.

The mean value of waiting time for all samples is 10.3 s, with a standard deviation of 19.4 s. For all red-phase arriving pedestrians, the mean waiting time for violators was 6.45 s (SD: 13.86), while the mean value of compliant ones was 33.01 s (SD: 26.03). The maximum waiting time was 166.64 s while the minimum was 0 s which indicates pedestrians who crossed the intersection without waiting.

7.3.1.1 Kaplan-Meier Analysis for Waiting Time

Pedestrians waiting at an intersection either wait for the green or cross the street during the red-signal phase. The choice of crossing in the red-signal phase is considered risk-taking behaviour. In the study, pedestrians who took the risk and crossed in the red signal phase were considered for the Kaplan-Meier curve estimation process. In contrast, pedestrians who waited and crossed in the green-signal phase were censored. The KM curve was estimated using Stata 17 software. The censored observations were coded as "0", whose actual waiting time was unknown, whereas pedestrians who crossed taking risks were coded as 1. The KM plot shows the waiting time interval on the x-axis and survival probabilities on the y-axis, as shown in Figure 7.1. The estimated probabilities for corresponding waiting time follow a similar trend as observed in studies conducted by Dhoke et al. (2021), Guo et al. (2012), and Tiwari et al. (2007). At the beginning of the waiting time, the curve declined very fast. The trend of the survival function illustrates that 49.5% of pedestrians crossed immediately after arrival (within 0-3 seconds) by violating the signal. This shows that certain pedestrians crossed the street immediately if there were a minimum number of approaching traffic or the violation tendency was predetermined. The violation percentage gradually increased by an additional 25.5% with an additional waiting time of 33.705 seconds. After waiting for 50 seconds, the survival function becomes stable as, after long waiting, pedestrians were no longer eager to take the risk and were willing to cross in green phase time.

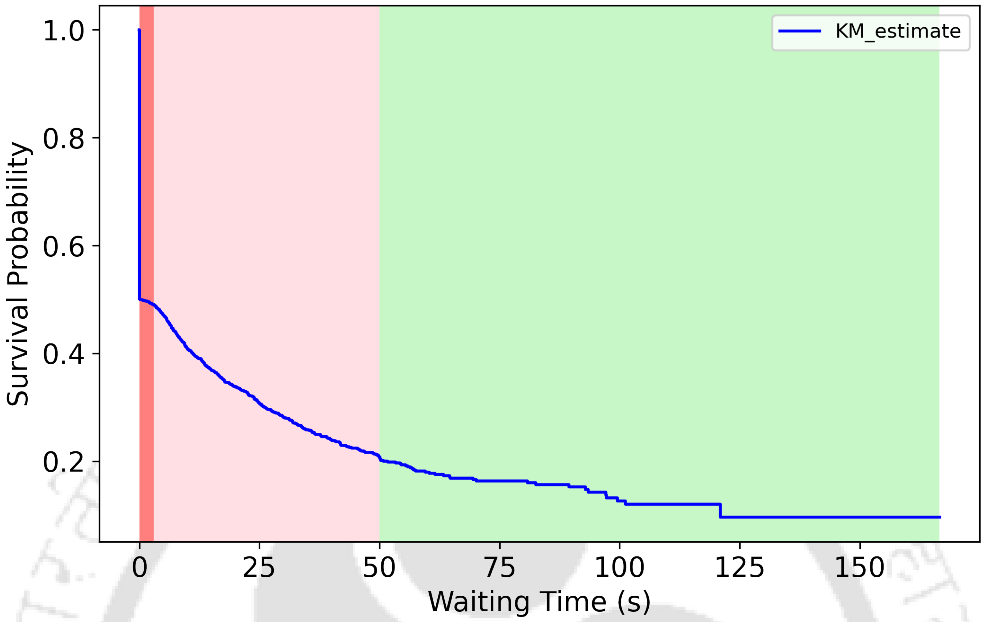


FIGURE 7.1: Kaplan-Maier Curve for the waiting time of pedestrians

7.3.1.2 Cox Proportional Hazard Model

To estimate the impact of different covariates on pedestrian road crossing behaviour, a semi-parametric Cox-hazard model was fitted using the 'lifelines' package in Python 3.6 (Cameron et al., 2019). It can be observed from Table 7.1 that 'luggage', 'crossing type', 'crossing pace', 'group talking', 'glance behaviour', 'cycle time', and the number of 'through and turning traffic per cycle' had a significant impact on the pedestrian signal violation. Pedestrian 'demographic' factors such as gender and age have no significant influence on the pedestrian red-light violation. The model estimates showed that the 'cycle length' at which a pedestrian arrived at the intersection, the 'direction of arrival' (oblique), and 'during crossing look at traffic' are the top positive factors that increase the signal violation tendency.

Similarly, variables such as 'before crossing look at traffic' and 'group talking' distraction are the topmost significant variables that reduce signal violation tendency.

The estimated Concordance Index (CI) value is 0.808, indicating that the overall model has good classification ability (good ability to distinguish between red and green crossing). The loglikelihood ratio test provided a significant p-value that indicated an overall good model fit compared to a null model (intercept-only model).

Table 7.1: Cox-based signal violation model estimates

Features	**Variable name**	**Coefficient** (β)	**HR** (e^{β})	**Standard error**	**P-value**
Demographic	Gender: Female	0.093	1.098	0.059	0.111
	Age Group				
	18-29	-0.056	0.945	0.103	0.583
	30-45	-0.138	0.871	0.099	0.165
	46-60	-0.11	0.895	0.108	0.307
	>60	-0.056	0.945	0.142	0.693
State of Crossing	Luggage: Yes **	0.149	1.161	0.059	0.012
	Crossing path: Oblique **	0.237	1.268	0.070	¡0.001
	Crossing pace: Hurried **	0.204	1.226	0.090	0.024
	Distraction				
	Mobile talking	-0.127	0.881	0.130	0.329
	Text	-0.163	0.849	0.282	0.562
	Headphones	0.025	1.025	0.215	0.908
	E/D/S	-0.088	0.915	0.187	0.637
	Group talking **	-0.181	0.834	0.091	0.048
	Holding phone *	-0.203	0.816	0.104	0.051
	Other **	-0.236	0.79	0.119	0.048
Cautionary Behaviour	Before crossing looking at traffic: Yes **	-0.890	0.410	0.061	<0.001
	During crossing looking at traffic: Yes **	0.208	1.231	0.080	0.010
	Looking frequency **	0.169	1.184	0.030	<0.001
Signal and Traffic Characteristics	***Signal Cycle***				
	>100-150 s **	1.168	3.215	0.101	<0.001
	>150 s **	1.660	5.257	0.107	<0.001
	Through traffic per cycle **	-0.008	0.991	0.001	<0.001
	Turning traffic per cycle **	-0.013	0.987	0.002	<0.001

Model fit statistics
Concordance = 0.808

Likelihood ratio test
Likelihood ratio test = 1056, on 22 df, p-value = <0.001

Reference levels
Gender: Male; Age: <18; Luggage: No; Crossing path: Straight; Crossing type: Normal; Distraction: No distraction; Before and during traffic look: No; Cycle time category: 0-100 s.

Note:

E/D/S: Eating/Drinking/Smoking.

****** Denotes variable significance at 95% confidence level;

***** Denotes variable significance at 90% confidence level.

Check for Proportional Hazard Assumption

There are two ways to test the proportional hazard assumption: one graphical approach and one goodness-of-fit test.

Both of them test the hazard ratio, which compares two specifications of the covariates (defined as $X^\star$ and X) can be expressed as:

$$HR = exp\left(\sum_{j=1}^{p} \beta_j (X_j^* - X_j)\right) \tag{7.17}$$

Where $X^\star = (X_1^\star, X_2^\star, \ldots, X_j^\star)$ and $X = (X_1, X_2, \ldots, X_j)$, and proportionality of hazard assumption indicates this quantity is constant over time. This indicates that the hazard of one individual is proportional to the hazard of another individual, where the proportionality constant is independent of time.

The graphical approach of the proportionality test involved plotting and comparing using log-log survival curves over different categories of variables. Another graphical option could be to use the Schoenfeld residuals to examine the model fit and detect outlying covariate values (Schoenfeld, 1982). The second approach for assessing the PH assumption involves the goodness of fit test (Grambsch and Therneau, 1994). This tests the correlation between the Schoenfeld residuals and survival time.

In the current study, the developed Cox-PH model has been tested based on a goodness-of-fit test using the R's survival package. The test results are presented in Table 7.2.

The results showed that the test was statistically significant for the 'looking behaviour', the 'crossing pace' and the 'turning traffic'. Additionally, the global test was also statistically significant. Therefore, the proportional hazard assumption was

Table 7.2: Goodness-of-fit test estimates

Variables	χ^2	DF	P-value
Gender	0.221	1	0.638
Age	6.35	4	0.174
Carrying luggage	0.668	1	0.414
Before crossing look at traffic **	5.628	1	0.018
During crossing look at traffic **	12.31	1	<0.001
During crossing look frequency **	13.246	1	<0.001
Crossing path	1.272	1	0.259
Crossing pace **	7.258	1	0.007
Distraction	9.719	7	0.205
Cycle time **	6.115	2	0.047
Through traffic per cycle	0.144	1	0.704
Turning traffic per cycle **	7.776	1	0.005
GLOBAL **	63.037	22	<0.001

Note:

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

violated. Thus, the hazard for the significant covariates was not constant over time. Due to the non-parallel nature of the curves, further investigation is needed when analysing data through a particular distribution.

7.3.1.3 Parametric AFT Model

In case the proportional hazard assumptions were violated, the result of cox proportional hazard model estimates no longer provide valid estimates. Therefore, an alternative option would be estimating the AFT models. In the current study, we fitted various AFT models for better interpretability of the results. The AFT model provides estimates in the form of acceleration factor, which provides the survival rate due to changes in the covariates. In the current study, we fitted and compared Weibull AFT, Exponential AFT, Log-Normal AFT, and Log-Logistic AFT distributions. The impact of covariates was enumerated based on the estimates, i.e., acceleration factors and their coefficient values which provided the change in duration because of the variations in covariates. The percentage change in survival due to change in covariates can be estimated using the following equation:

$$\%Change = [e^{\beta} - 1] * 100\% \tag{7.18}$$

The positive coefficient value shows an increase in survival time with a unit increase in parameters of waiting duration. To select the best AFT model, the fitted models were compared with three goodness of fit criteria, that is, likelihood ratio, AIC and BIC value shown in Table 7.3. The result implied that the Log-Normal distribution best fits the waiting duration data. The model estimates of the fitted Log-Normal AFT model are presented in Table 7.4. The estimated Concordance Index (CI) value is 0.808, indicating that the overall model has good classification ability that can distinguish between red-light and green-light crossing.

TABLE 7.3: Goodness of fit of different models

Distribution	**DOF**	**Log-Likelihood**	**AIC**	**BIC**
Weibull AFT	24	-4370.752	8789.505	8924.972
Exponential AFT	23	-6451.353	12948.710	13078.530
Log-Normal AFT	24	-4286.698	8621.396	8756.863
Log-Logistic AFT	24	-4318.229	8684.458	8819.925

The variables ranking based on the log(acceleration factor) illustrated in Figure 7.2. It can be observed that 'before crossing look at traffic' and 'group talking' were the top influencing covariates that positively influenced the survival time. Similarly, 'cycle time' and 'crossing path' were the top covariates that decreased survival time. For better interpretability of the covariates, survival functions of the fitted Log-Normal distribution have been plotted using Stata 17 software and illustrated in Figure 7.3.

Table 7.4: Estimated results of log-normal AFT signal violation model

Features	Variable name	Coefficient (β)	Acceleration factor	Standard error	P-value
Demographic	Gender: Female	-0.208	0.812	0.166	0.209
	Age Group				
	18-29	0.242	1.274	0.294	0.410
	30-45	0.430	1.537	0.284	0.130
	46-60	0.384	1.469	0.304	0.206
	>60	0.056	1.058	0.404	0.889
State of Crossing	Luggage: Yes **	-0.478	0.620	0.160	0.003
	Crossing path: Oblique **	-0.923	0.397	0.200	<0.001
	Crossing pace: Hurried **	-0.747	0.473	0.266	0.005
	Distraction				
	Mobile talking	0.261	1.299	0.360	0.468
	Text	0.560	1.749	0.705	0.428
	Headphones	-0.016	0.984	0.613	0.979
	E/D/S	0.639	1.895	0.537	0.234
	Group talking **	0.735	2.086	0.254	0.004
	Holding phone *	0.568	1.765	0.290	0.051
	Other **	0.845	2.327	0.343	0.014
Cautionary Behaviour	Before crossing look at traffic: Yes **	2.511	12.315	0.172	<0.001
	During crossing look at traffic: Yes **	-0.615	0.540	0.221	0.005
	Looking frequency **	-0.501	0.582	0.090	<0.001
Signal and Traffic Characteristics	***Signal Cycle***				
	>100-150 s **	-3.397	0.033	0.240	<0.001
	>150 s **	-4.770	0.008	0.258	<0.001
	Through traffic per cycle **	0.022	1.023	0.001	<0.001
	Turning traffic per cycle **	0.032	1.032	0.006	<0.001

Model fit statistics
Concordance = 0.808

Reference levels
Gender: Male; Age: <18; Luggage: No; Crossing path: Straight; Crossing type: Normal; Distraction: No distraction; Before and during traffic look: No; Cycle time category: 0-100 s.

Note:

E/D/S: Eating/Drinking/Smoking.

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

7.4 Discussion

A systematic discussion of obtained results is presented in the subsequent sections. In the model estimate table (Table 7.4), a positive value of the coefficient indicates an increase in waiting duration (reduction in violation tendency). Similarly, a negative coefficient indicates a decrease in the waiting duration (increase in violation

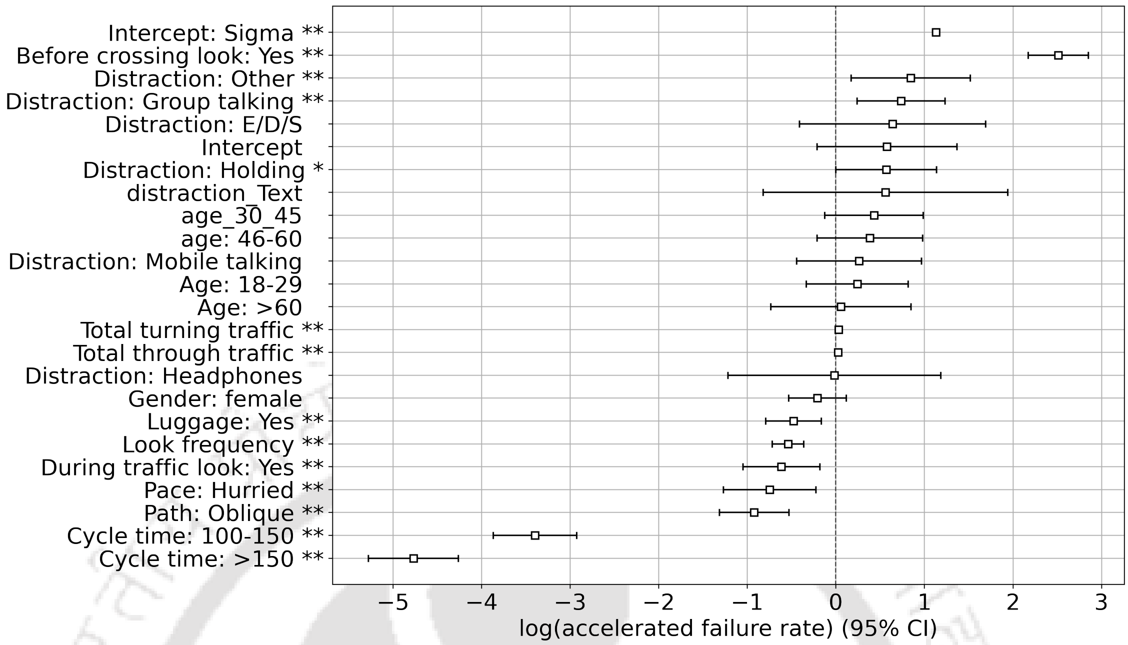


FIGURE 7.2: Log-Normal AFT model variable ranking based on log(accelerated failure rate).

Note:

** Denotes variable significance at 95% confidence level; * Denotes variable significance at 90% confidence level.

tendency).

7.4.1 Demographic Characteristics

In the present study, the gender and age of pedestrians did not emerge as essential factors in signal violation (Table 7.4). The reason could be because in developing countries like India, where traffic rules are lenient, pedestrians, irrespective of gender and age accustomed to risky road crossing behaviour. This can also be corroborated by the fact that West Bengal (Indian state), where the selected study locations are situated, possessed the highest number of pedestrian fatalities during the study year. The results align with the past study conducted by Ren et al. (2011), who also reported no significant influence of gender in road crossing choice. Further, Rosenbloom (2009) studied the impact of single vs group behaviour in signal violation and observed no significant age difference in road crossing behaviour. However, other observational studies (investigating signal violation behaviour) revealed a significant impact of age on the signal violation (Brosseau et al., 2013; Cooper et al., 2012; Praveen Kumar N. et al., 2018; Nasar and Troyer, 2013), especially among the young adult age group (18-35 years).

7.4.2 State of Crossing

According to the model estimates, the variable 'carrying luggage' (reference: not carrying luggage) provided a negative coefficient estimate of -0.478 (Table 7.4), indicating a higher signal violation likelihood among pedestrians carrying luggage. It was further observed that pedestrians with luggage were found to be more impatient and committed 38% more violations (lower waiting endurance/ survival time) than those without luggage (Table 7.5). The possible reason could be that those carrying luggage could be in a hurry to reach their destination, thus violating the signal. A similar finding was also observed by Mukherjee and Mitra (2020), in Kolkata city, where pedestrians were more likely to violate the signal while carrying overhead loads. Similarly, pedestrians who started crossing sneaking from the side of the crosswalk marking or crossed following an oblique path were 60.27% more likely to violate the signal than those who crossed perpendicularly. Furthermore, pedestrians in a hurry had less patience thus, waited less.

The distraction variable estimates showed that mobile phone talking, texting, headphones use, and eating/drinking/smoking had no significant influence on the red-light violation behaviour. The positive coefficient value corresponding to 'group talking' indicated that a pedestrian in a group conversation waits 108.55% longer compared to a pedestrian without distraction. The current observation contradicts the finding reported by Thompson et al. (2013), who reported that pedestrians are 1.69 times more likely to exhibit unsafe crossing behaviour when conversing in a group. The possible reason could be that pedestrians conversing in a group might frequently look towards each other rather than paying attention to the traffic, increasing the crossing risk during the red-signal phase. Thus, to be on the safe side, pedestrians in a group try to avoid taking unnecessary risks. Another possible reason behind long waiting behaviour could be that in a social context (in a group setting), pedestrians usually fear all forms of social criticism by other group members if the subject pedestrian is going against the social norm (group ethics), i.e., crossing against the red-light phase (Pele et al., 2017; Rosenbloom, 2009).

Table 7.5: Percentage change in survival for significant variables.

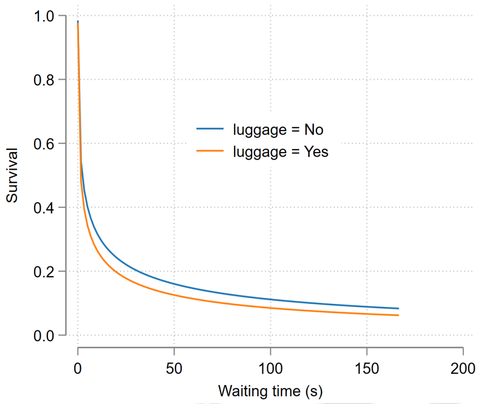
Variable name	**Coefficient** (β)	**% Change in survival** $(e^{\beta}-1)\times 100\%$
Luggage: Yes **	-0.478	-38.000
Crossing path: Oblique **	-0.923	-60.267
Crossing pace: Hurried **	-0.747	-52.620
Distraction: Group talking **	0.735	108.548
Distraction: Holding phone *	0.568	76.473
Distraction: Other **	0.845	132.790
Before crossing looking at traffic: Yes **	2.511	1131.720
During crossing looking at traffic: Yes **	-0.615	-45.936
Looking frequency **	-0.501	-39.407
Cycle time: >100-150 s **	-3.397	-96.653
Cycle time: >150 s **	-4.770	-99.152
Through traffic per cycle **	0.022	2.224
Turning traffic per cycle **	0.032	3.252

Note:

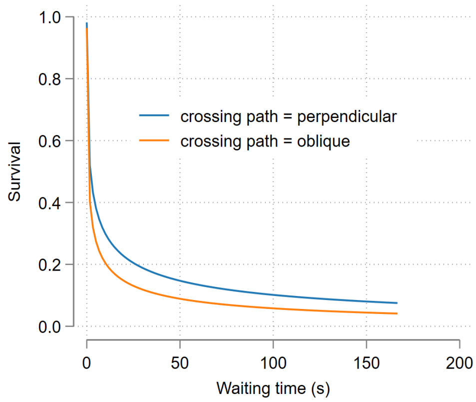
The +ve sign indicates an increase in survival time, while -ve sign indicates a decrease in survival time.

** Denotes variable significance at 95% confidence level;

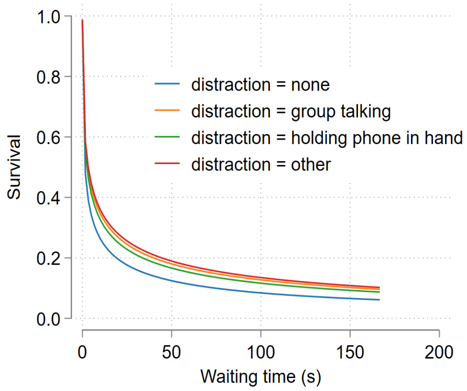
* Denotes variable significance at 90% confidence level.



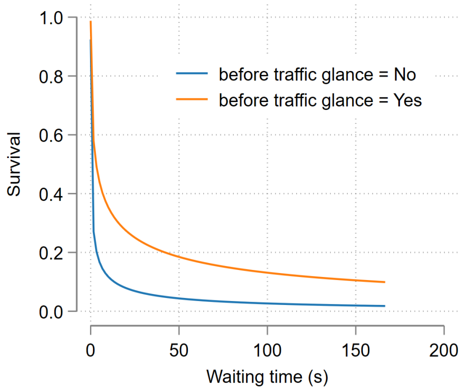
(A) Luggage categories



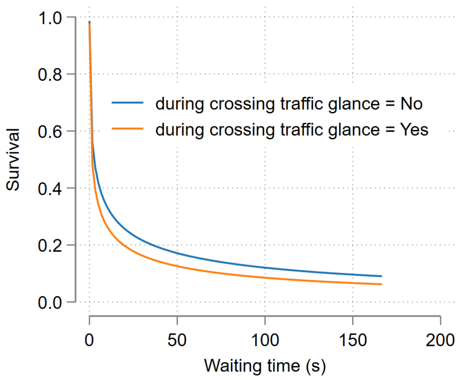
(B) Path categories



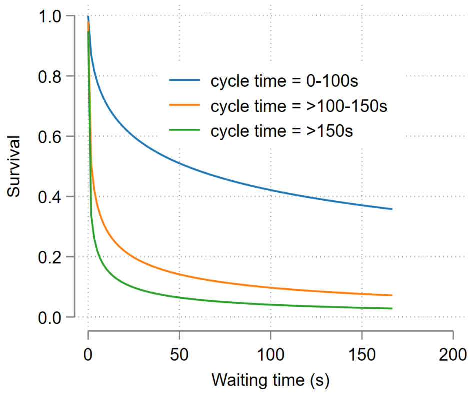
(C) Distraction categories



(D) Before crossing glance categories



(E) During crossing glance categories



(F) Cycle time categories

Figure 7.3: Log-Normal AFT model's survival function comparison for different covariates

7.4.3 Cautionary Behaviour

Pedestrian cautionary or glance behaviour is an essential factor that provides important information to pedestrians, which helps them to plan a better and safer crossing decision. The 'before crossing look' (glance) at the traffic is the foremost significant predictor that helps in taking the crossing decision. The Table 7.4 model estimate showed a positive coefficient (2.511) for 'before crossing look at traffic' variable, indicating pedestrians who looked at traffic before finally deciding to cross were more likely to wait (survival wait time is 1131.72% more) than those who did not look at all, which is a significant finding. The reason for such behaviour could be that, before starting to cross, pedestrians try to assess the risk associated with the red-light crossing. The 'before-crossing look at traffic' behaviour helps pedestrians to quantify the risk. So, those who looked at traffic got familiarised with the risk associated with crossing, and in the context of the current study, this perceived risk might have been greater than the pedestrians' threshold risk-taking capacity; thus, they avoided crossing during the red-signal phase.

When pedestrians decide to cross against the red-light phase, the looking at traffic behaviour entirely changes compared to the before-crossing look. During crossing against the red light, pedestrians start observing traffic more frequently. The reason behind that during the signal violation, pedestrians are at constant risk of traffic exposure. In addition, they needed to accept a minimum safe gap between vehicles. For that reason, one has to constantly keep track of the oncoming vehicle's position, which requires more frequent looking. The current finding is consistent with the research conducted by Thompson et al. (2013), which reported that pedestrians display more caution (look left and right) when crossing against the traffic signal than those who follow the traffic signal.

7.4.4 Traffic and Signal Characteristics

The impact of the varying phase length can also be observed from the 'cycle length' variable. The model estimate showed that in comparison to cycle length of 0 to 100 seconds, pedestrians who arrived at the signal cycle category of 100-150 seconds (or >150 seconds) were 96.653% (or 99.152%) less likely to wait for safe crossing. Additionally, the model estimate showed that traffic count per cycle significantly affected signal violation behaviour. As the number of vehicles per cycle increased, pedestrians were less likely to take the violation risk. The availability of safe gaps decreases with the increase in the number of vehicles, discouraging pedestrians from accepting such unsafe gaps to violate the signal.

7.5 Practical Implications from the Results

- The study result revealed that the hazard-based duration approach could be a very useful technique for understanding the signal violation behaviour over a given duration of the waiting period.
- Log-Normal distribution best represents the waiting duration of pedestrians at intersection crosswalks in Kolkata city, which could be very useful information for planners and traffic engineers for designing future infrastructure.
- The study results revealed that the high danger rate is shown by short waiting time (less than 3 seconds), where 49.5% of pedestrians crossed the road immediately after red-phase arrival, indicating a pre-determined crossing decision or familiarity with the intersection. Intersections with many early crossers should be given special attention, and interventions must be implemented to reduce such unsafe crossing behaviour.

- The pedestrian endurance towards waiting was reflected in the span of 36.705 seconds, over which 75% of the pedestrian violated the signal. The present finding points to the deficiencies in signal design and infrastructure planning.

- The pedestrians carrying luggage were observed to be more likely to violate traffic signal rules (i.e., displayed lower endurance for waiting). Hence, intersections with a high volume of such pedestrians require additional attention.

- Observation on distracted road crossing behaviour revealed that 'group talking' social distraction decreases signal violation tendency and increases the waiting endurance for the green light. Although the impact of distraction in the current study is not unsafe in the observed intersections, still wherever this type of behaviour is found to be unsafe during road crossing on those intersections, innovative interventions should be applied.

- A significant finding is that pedestrian glance is an essential predictor of signal violation. The study results revealed that pedestrians who looked at traffic before initiating crossing displayed higher endurance for waiting (waited longer). Thus, educating pedestrians about the importance of traffic glances in road crossing could help them cross safely.

- The model outcome also suggested that longer cycle length increases the probability of signal violation, as the pedestrian has to wait longer for a safe crossing opportunity. This is an important finding, capturing broader planning, design and enforcement issues influencing pedestrian risk-taking behaviour. Redesigning the pedestrian signal setting with updated information could help increase pedestrian signal compliance at busy intersections.

7.6 Summary

In this section, pedestrian waiting behaviour was examined using 2089 (red-light phase arriving) pedestrian samples collected from eight signalised intersection crosswalks in Kolkata, India. The study began with an initial estimation of a Kaplan-Meier curve (non-parametric) to estimate the survival probabilities (waiting), which provided an overview of how long a certain proportion of observed pedestrians will wait in the red-light phase. Though the KM estimate is informative, it still can't be used for checking the impact of different covariates. To cope with the problem, a semi-parametric Cox Proportional Hazard model was used to understand the factors influencing pedestrian waiting duration during the red-light phase. However, the proportional hazard assumption was not satisfied. Therefore, the parametric Accelerated Failure Time (AFT) model was used to understand the covariates affecting the waiting duration. Based on the results, several practical interventions were proposed. The next chapter discusses pedestrian-vehicle interaction at signalised intersection crosswalks using the safety margin approach.

Chapter 8

Safety Margin Modelling at Signalised Intersection Crosswalks

8.1 General

The pedestrian risky road crossing behaviour at intersection crosswalks depends on the pedestrian gap acceptance mechanism. In the gap acceptance process, the pedestrian maintains a gap margin value with approaching traffic to complete the crossing manoeuvre safely. Accepted gap size, time of crossing, and pedestrian and traffic characteristics significantly influence the margin value maintained by pedestrians. The safety margin defines as the gap between the time a pedestrian cross before a vehicle and the time it arrives at the crossing point. Pedestrian Safety Margin (PSM) is a kind of dummy conflicting measurement between pedestrians and vehicles, and it plays a vital role in crossing safety. Hence, the current study evaluates pedestrian safety in terms of PSM. The current chapter presents PSM models for intersection crosswalks. The important factors, which influence the PSM while crossing at intersection crosswalks under mixed traffic conditions, are also discussed in this chapter.

8.2 Methodology

In the context of modelling, usually researcher attempt to understand the relationship between a dependent (PSM) and independent variables. To understand the influence of various explanatory variables on the PSM, an attempt has been made to understand pedestrian safety during road crossing at intersection crosswalks. As PSM is a continuous variable, Multiple Linear Regression (MLR) is the best-suited approach to model and explore the relationship.

8.2.1 Model Formulation for PSM

In the MLR model, PSM is the dependent (response) variable and the site, demographic characteristics, pedestrian characteristics, exposure, state of crossing, glance and traffic characteristics are considered explanatory variables. In this study, the MLR model has been utilised to understand the safety margin value maintained by a pedestrian with approaching traffic while crossing through the intersection crosswalk. The MLR model framework is given by Eq. (8.1):

$$PSM(\hat{Y}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n + \epsilon_i \tag{8.1}$$

Where PSM = Pedestrian Safety Margin (response variable); X_{i-n} = explanatory variable ; β_{i-n} = estimated parameters of the model; β_0 = y-intercept of regression line; ϵ_i = prediction error for the i_{th} pedestrian.

The standard MLR model assumes that the effects of variables are fixed across observation sites. However, the individual's behaviour is influenced by the group to which they belong (here the site observed). Ignoring the between-group variance (among sites) leads to biased parameter estimates. Therefore, the random intercepts are considered to capture the heterogeneity among the eight sites (two-way crosswalks). The framework of the random intercept linear regression model is given by Eq. (8.2):

$$PSM(\hat{Y}_{ij}) = \beta_0 + \beta X_{ij} + u_j + \epsilon_{ij} \tag{8.2}$$

where X_{ij} is a vector of covariates for i^{th} individual in group j, β is the coefficient, β_0 is the constant, u_j and ϵ_{ij} are the group specific random intercept and individual level error term, respectively, $\beta_0 + u_j$ represents intercept for group j.

8.3 Goodness of Fit Measure

For MLR fixed effect model, the commonly used goodness of fit measures is the F-test (overall significance of model) and coefficient of determination (R-square). Similarly, for a random intercept regression model, the commonly used goodness of fit measure is the Wald chi-squared test. It is also essential to check for the significance of explanatory variables to include them in the model. In the current study, F-test has been used to test the goodness of fit of the one-way PSM model, while the Wald Chi-squared test was used for the two-way PSM model, as discussed below.

8.3.1 Overall Significance of the MLR Model

The F-test is used to test the overall significance of the fixed-effects MLR model, indicating whether the coefficients are statistically different from zero. The hypothesis being tested are as follows:

- H_0: The null hypothesis is that there is no significant effect of selected independent variables on the outcome variable (PMS).
- H_1: The alternative hypothesis is that there is a significant impact of independent variables on the outcome variable (PMS).

If the estimated F-value is higher than the critical value for p-value <0.05, then there is evidence of a strong linear relationship between the outcome and explanatory variables.

The F-statistics value is calculated using the following equation:

$$F = \frac{\frac{SS_R}{q}}{\frac{SS_E}{N-q-1}} \tag{8.3}$$

Where: SS_R = Regression Sum of Squares; SS_E = Residual Sum of Squares; q = degree of freedom for regression analysis; N-q-1 = degree of freedom for residual analysis.

8.3.2 R-Square

Subsequently, how well the predictors and response variables fitted in the model was expressed using different evaluation metrics. One of the most common evaluation metrics in MLR is $R-square$, which ranges from 0 to 1. A $R-square$ of 0 indicates that the estimated model does not account for the variability of selected explanatory variables. In contrast, the $R-square$ value of 1 indicates a perfect prediction of PSM. In general, $R-square$ indicates the percentage of variance explained by the estimated model.

$$R-square\ (R^2) = 1 - \frac{SS_E}{SS_T} = \frac{SS_R}{SS_T} = \frac{\sum(\hat{Y}-\bar{Y})^2}{\sum(Y-\bar{Y})^2} \tag{8.4}$$

Where: SS_E = Residual sum of squares; SS_T = Total sum squares; SS_R = Regression sum of square; $\hat{Y}$ = Predicted value; $\bar{Y}$ = Mean value.

However, adding more explanatory variables to the model may increase the $R-square$ value even if the variable has an insignificant contribution. To overcome this problem, in place of $R-square$, a $adjusted\, R-square$ metric used. The $adjusted\ R-square$ accounts for the changes in the additional variable and penalised addition of insignificant variables by reducing the $R-square$ value.

In the model development process, multicollinearity among predictors creates unreliable estimates. To overcome this problem, researchers conduct either a pairwise correlation and eliminate predictors having a correlation value beyond the threshold limit (rule of thumb, $r \geq 0.4$) or estimate an index called Variance Inflation Factor (VIF). The VIF can be estimated as follows:

$$VIF_i = \frac{1}{1-R_i^2} \tag{8.5}$$

Where: n = number of explanatory variable, and R_i^2 = unadjusted multiple regression coefficient of the i^{th} variable with the remaining (n-1) variables. The VIF value could be interpreted as:

- VIF score 0-5: no evidence of multicollinearity

- VIF score 5-10: moderate multicollinearity
- VIF score >10: serious multicollinearity

8.3.3 Wald Chi-squared Test

The overall significance of the random intercept regression model was tested with Wald Chi-squared (χ^2) test. The test estimates help verify whether the regression model coefficients differ from zero. The hypothesis tests the following:

- **Null Hypothesis** (H_0): There is no significant effect of any explanatory variables on PSM, and the regression coefficients are zero.
- **Alternative Hypothesis** (H_1): There is a significant effect of at least one explanatory variable on PSM, and at least one regression coefficient is different from zero.

8.4 PSM Significant Variable Selection

In a regression model, adding a large number of variables improves the R-square but increases the model complexity, leading to model overfitting. Hence, identifying influential variables that significantly contribute to PSM prediction is very much essential. The common statistical method for variable selection is ANOVA, which assumes that the dependent (outcome) variable is normally distributed. Hence, whether the dependent variable follows a normal distribution needs to be verified using hypothesis testing. The most common test is the Shapiro–Wilk test for goodness of fit.

As two separate PSM models will be developed thus, PSM data for the one-way and two-way need to be verified separately for normally. The Shapiro–Wilk test hypothesis consists of the following:

- **Null Hypothesis** (H_0): The null hypothesis states that the variable is normally distributed.
- **Alternative Hypothesis** (H_1): The alternative hypothesis states that the variable is not normally distributed.

The Shapiro–Wilk test was performed with one-way (603 samples) and two-way (915 samples) safety margin data. The test statistics [one-way data W = 0.924, p-value = <0.001; two-way data (W) = 0.936, p-value = <0.001] revealed that both one-way and two-way safety margins does not follow a normal distribution. Hence, it does not satisfy the assumption for conducting an ANOVA test. An alternative approach is to use a non-parametric method for variable selection. Kruskal-Wallis analysis of variance test is a common non-parametric test used for this purpose.

The hypothesis being tested consists of the following:

- **Null Hypothesis** (H_0): The null hypothesis is that the samples come from populations with the same distribution.
- **Alternative Hypothesis** (H_1): The alternative hypothesis is that the samples do not come from populations with the same distribution.

The Table 8.1 and Table 8.2 revealed the test statistics for non-parametric ANOVA (Kruskal–Wallis) for one-way and two-way data, respectively. Test statistics revealed that for one-way data, 'gender', 'before crossing glance', 'group size', 'crosswalk

Table 8.1: Kruskal–Wallis test for safety margin (one-way crosswalks data).

Variable	χ^2	DF	P-value
Study location	14.251	2	<0.001**
Gender	0.019	1	0.8911
Age	11.091	4	0.026**
Before crossing traffic glance	0.000	1	0.9842
Group size	1.745	1	0.186
Crosswalk violation	1.468	1	0.226
Vehicle's position when safety margin taken	15.252	2	<0.001**
Type of vehicle with which safety margin has been taken	5.456	3	0.1413
Luggage	0.215	1	0.6429
Distraction	3.164	7	0.869
Speed category	8.580	4	0.072*
During crossing glance (traffic)	6.752	1	0.009**
During crossing glance (straight)	4.152	1	0.042**
During crossing glance (ground)	0.020	1	0.888

Note:

** Denotes variable significance at 95% confidence level;

* Denotes variable significance at 90% confidence level.

violation', 'vehicle type', 'carrying luggage', 'distraction' and 'during crossing glance (ground)' do not show any significant influence on safety margin.

Similarly, test statistics for two-way data revealed that pedestrian 'age', 'carrying luggage', and 'during crossing glance' (straight and ground) did not have a significant influence on safety margin.

8.4.1 Association Among Variables

In addition to variable selection, it is a common practice to estimate multicollinearity among predictors using a pairwise correlation. This is done to eliminate predictors having a correlation value beyond a threshold limit (rule of thumb, $r \geq 0.4$). In the present study, as the majority of the variables are of categorical type, thus instead

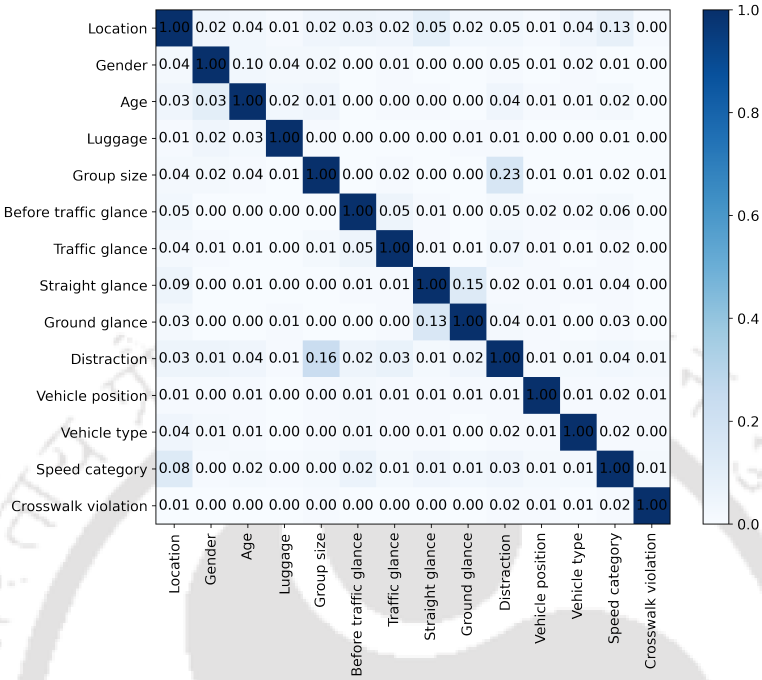
Table 8.2: Kruskal–Wallis test for safety margin (two-way crosswalks data).

Variable	χ^2	DF	P-value
Gender **	9.497	1	0.002
Age	3.542	4	0.471
Luggage	2.406	1	0.121
Group size **	17.854	1	<0.001
Before crossing traffic glance **	9.544	1	0.002
During crossing traffic glance **	10.693	1	0.001
During crossing glance (straight)	0.062	1	0.803
During crossing glance (towards ground)	0.241	1	0.624
Distraction **	20.256	7	0.005
Vehicle's position when safety margin taken *	4.849	2	0.088
Type of vehicle with which safety margin has been taken **	16.492	3	<0.001
Approaching vehicle's direction (straight/turning) **	37.038	2	<0.001
Speed category **	31.062	4	<0.001

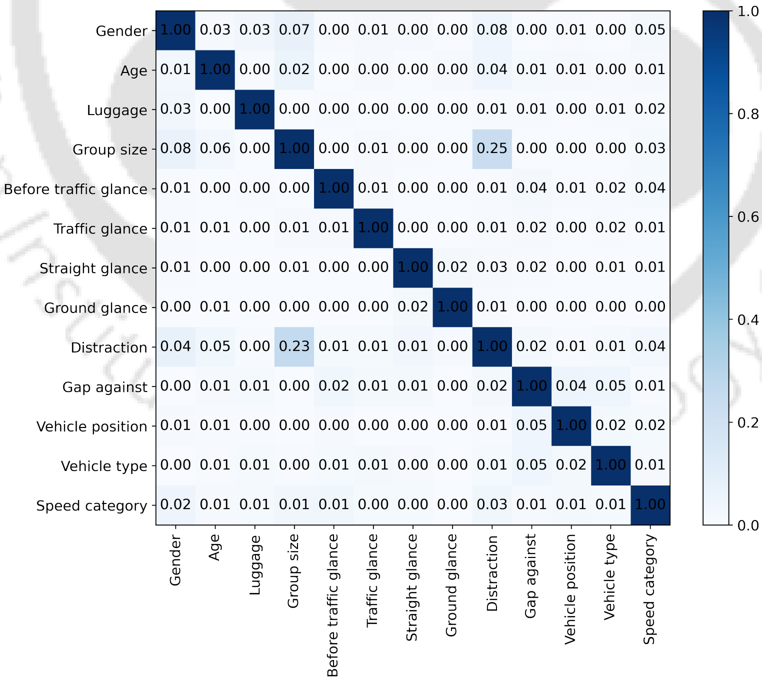
Note:

** Denotes variable significance at 95% confidence level;
* Denotes variable significance at 90% confidence level.

of estimating Pearson correlation coefficients, Theil's U (Theil, 1958, 1966) statistics for one-way and two-way variables were estimated and reported in Figure 8.1 (A) and 8.1 (B) respectively. Theil's U is an asymmetric measure of association [U(x, y) $\neq$ U(y, x), where U is Theil's U], which provides association strength for categorical variable pairs ranging from 0 to 1, where 0 indicates no association and 1 indicates perfect association. The values reported in Figure 8.1 (A) and 8.1 (B) revealed that the association strength among categorical variables is weak (Theil's U for one-way: $\leq$0.23; Theil's U for two-way: $\leq$0.25), indicating that the selected variables are independent and can be used in the model estimation process.



(A) Pairwise Theil's U association estimate among one-way covariates.



(B) Pairwise Theil's U association estimate among two-way covariates.

FIGURE 8.1: Pairwise association plots for categorical variables.

8.5 PSM Model Results

Two separate regression models have been estimated to measure the effects of explanatory variables (significant variables obtained from the non-parametric Kruskal–Wallis test) on PSM for one-way and two-way crosswalks. The fixed effect model estimates for the one-way crosswalk are presented in Table 8.3. Similarly, the fixed effect model estimates along with location-wise intercepts are presented in Table 8.4, and random-effects parameter estimates are presented in Table 8.5.

In addition to raw coefficients, semi-elasticities [**eydx(x)**] were computed using Stata 17 software for continuous and discrete variables ($\boldsymbol{x}$) using the Eq. (8.6) and Eq. (8.7) respectively.

$$eydx_{(continuous\ \ \mathbf{x})} = \frac{d(\ln \hat{y})}{dx} = \frac{dy}{dx} * \frac{1}{\hat{y}} = \frac{\hat{b}}{\hat{y}} \tag{8.6}$$

$$eydx_{(discrete\ \ \mathbf{x})} = E\{\ln(\hat{y})|x = 1\} - E\{\ln(\hat{y})|x = 0\} \tag{8.7}$$

The backward elimination process has been adopted for model fitting, where with each iteration, non-contributing variables were removed one by one until no improvement in model fitting was observed.

8.5.1 Overall Significance of the MLR Results

The overall significance of the one-way MLR model (one-way crosswalk data) is tested with $F - test$, and it has been estimated to check whether the regression coefficients are different from zero. The overall significance of the MLR model results is tabulated in Table 8.3. The calculated $F - values$ is 6.080 (for one-way model) with a significant $p-value$ (<0.001). It indicates that the one-way MLR model can predict the PSM with selected explanatory variables, and the alternative hypothesis has been accepted. The overall significance of the random intercept regression model results is presented in Table 8.4 (refer to next section). The test statistics [χ^2 (23) = 105.68] and a significant $p-value$ (<0.001) indicate that the model can predict the PSM with selected explanatory variables.

The one-way MLR model results are shown in Table 8.3. The MLR model's coefficient of determination (R-square) is 0.128. It means that there is only a 12.8% variation in PSM for the one-way model due to variation in the selected explanatory variables. The VIF values (mean VIF is 1.64) revealed that VIF values are within the acceptable limit (VIF $\leq$5) for the one-way model, which also reflects the independence of covariates.

A likelihood ratio test was conducted to compare the random intercept model with one-level ordinary linear regression [model without random intercept (u_j)]. The results revealed a significance p-value $<$ 0.05 (refer Table 8.5), indicating that the random intercept model is overall better compared to the fixed effect model.

8.5.2 Variance Partitioning Coefficient

To understand the variability explained by the random intercept model, the Variance Partitioning Coefficient (ρ) was estimated using Eq. (8.8). The random-effects parameter (here the location) estimates revealed that it explained an additional 12.83% (ρ = 0.128) of variability, which was not accounted for by the fixed effect model.

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \tag{8.8}$$

where σ_e^2 = residual variance at level 1 (2.085), σ_u^2 = variance at level 2 (0.307) and $\sigma_u^2 + \sigma_e^2$ = total residual variance (2.392).

8.6 Discussion on PSM Model Results

From the developed model results of one-way crosswalks (refer Table 8.3), revealed that 'gender', 'age', 'before crossing traffic glance', 'group size', 'luggage', 'distraction' and 'during crossing glance' (straight and ground) do not show any significant effect on PSM model. Similarly, results from the two-way crosswalk model (refer Table 8.4) revealed that 'gender', 'age', 'before crossing glance', 'luggage', 'vehicle position', and 'during crossing look' at straight and ground do not show any significant influence. The various significant variables of one-way and two-way models are described as follows:

8.6.1 Effect of Demographic Characteristics

In both one-way and two-way models, pedestrian demographic such as age and gender do not show any significant influence. One reason could be that in developing countries like India, where traffic rules are lenient or at times ignored deliberately, pedestrians, irrespective of gender and age, are accustomed to risky road crossing behaviour. A similar observation was also found in the works of Onelcin and Alver (2015), who studied illegal crossing behaviour at signalised intersections using gap acceptance. However, past studies conducted on mid-block sections reported a significant influence of age and gender on safety margin (Chaudhari et al., 2020).

TABLE 8.3: Pedestrian safety margin model results for one-way crosswalks

Features	Variable name	Coefficient (β)	Std. err.	p-value	ey/dx	VIF
	Constant **	5.764	0.505	<0.001	-	-
Site	Site 2 (BBG) **	-0.614	0.253	0.015	-0.163	1.82
	Site 3 (GPO) **	-1.619	0.288	<0.001	-0.520	3.20
Pedestrian behavioural characteristics	Crossing speed: >1-1.2 (m/s) **	-0.401	0.198	0.044	-0.119	1.57
	Crossing speed: >1.2-1.4 (m/s) **	-0.638	0.214	0.003	-0.197	1.54
	Crossing speed: >1.4-1.5 (m/s) **	-0.970	0.257	<0.001	-0.318	1.27
	Crossing speed: >1.5 (m/s) **	-0.940	0.235	<0.001	-0.306	1.91
Exposure	Vehicles position: Middle	0.000	0.157	1.000	0.000	1.36
	Vehicles position: End **	-0.705	0.207	0.001	-0.252	1.34
	Vehicle type: Car **	0.401	0.167	0.017	0.140	1.39
	Vehicle type: LCV	0.477	0.308	0.122	0.164	1.28
	Vehicle type Bus **	0.779	0.298	0.009	0.255	1.26
Cautionary behaviour	Looked straight during crossing: yes **	0.340	0.157	0.031	0.115	1.15
	Looked at traffic during crossing: yes **	-0.422	0.169	0.013	-0.133	1.07
Traffic characteristics	Through traffic **	-0.042	0.009	<0.001	-0.014	2.41

Model fit statistics
Multiple R-squared = 0.128
AIC: 2354.543; BIC: 2420.571
F(14, 588): 6.080 , P-value: <0.001

Reference levels
site = site1; crossing speed = ≤1; vehicle position = initial lane; vehicle type = two-wheeler; looked at traffic = no; looked straight = no.

Note:

****** Denotes variable significance at 95% confidence level;
***** Denotes variable significance at 90% confidence level.

Table 8.4: Pedestrian Safety Margin model summary for two-way crosswalks.

Features	**Variable name**	**Coefficient (β)**	**Std. err.**	**z**	**p-value**	**ey/dx**
Demographic	Gender (Female)	-0.084	0.116	-0.724	0.469	-0.028
Pedestrian behavioural characteristics	Group: Two or more **	0.337	0.167	2.024	0.043	0.110
	Crossing speed: >1-1.2 (m/s) *	-0.291	0.166	-1.750	0.080	-0.085
	Crossing speed: >1.2-1.4 (m/s) **	-0.610	0.182	-3.363	0.001	-0.187
	Crossing speed: >1.4-1.5 (m/s) **	-0.532	0.206	-2.575	0.010	-0.161
	Crossing speed: >1.5 (m/s) **	-0.910	0.171	-5.312	<0.001	-0.295
Exposure	Vehicles position: Middle	0.129	0.119	1.082	0.279	0.044
	Vehicles position: End	0.018	0.161	0.109	0.913	0.006
	Vehicle type: Car *	0.236	0.123	1.922	0.055	0.082
	Vehicle type: LCV	0.242	0.185	1.309	0.190	0.084
	Vehicle type: Bus **	0.579	0.217	2.665	0.008	0.189
	Approching vehicle's direction: Left **	0.379	0.130	2.908	0.004	0.129
	Approching vehicle's direction: Right **	0.424	0.163	2.597	0.009	0.144
State of crossing	Distraction: Mobile talking	-0.196	0.234	-0.836	0.403	-0.068
	Distraction: Texting	-0.260	0.518	-0.503	0.615	-0.091
	Distraction: Headphones *	-0.659	0.385	-1.713	0.087	-0.251
	Distraction: Eating/drinking/smoking	0.452	0.443	1.021	0.307	0.140
	Distraction: Group talking	-0.020	0.215	-0.094	0.925	-0.007
	Distraction: Holding phone in hand	-0.011	0.212	-0.051	0.959	-0.004
	Distraction: Others	-0.026	0.279	-0.092	0.927	-0.009
Cautionary behaviour	Before crossing traffic glance: yes	-0.197	0.128	-1.544	0.123	-0.065
	During crossing traffic glance: yes **	-0.402	0.136	-2.952	0.003	-0.129
Traffic characteristics	Traffic per cycle **	-0.004	0.001	-3.420	0.001	-0.001
Intercepts	Central	4.879	na	na	na	na
	Chandni Chowk	3.405	na	na	na	na
	Deshapriya Park	4.447	na	na	na	na
	Gariahat	3.749	na	na	na	na
	General Post Office	3.394	na	na	na	na
	Jadavpur	4.669	na	na	na	na
	Kalighat	3.723	na	na	na	na
	Shobhabazar	4.215	na	na	na	na

Model fit statistics
AIC: 3343.46 and BIC: 3468.75
Wald χ^2 (23) = 105.68, p = <0.001

Reference levels
Gender = Male; Group/Platoon size = Single; Crossing speed = $\leq$1; Vehicle position = Initial lane; Vehicle type = Two-wheeler; Approaching vehicle's direction = Straight; Distraction = None; Before crossing glance = No, and During crossing glance = No.

Note:

na: not applicable.

** Denotes variable significance at 95% confidence level;
* Denotes variable significance at 90% confidence level.

TABLE 8.5: Pedestrian Safety Margin model's random-effects parameters summary.

Random-effects parameters	Estimate	Std. err.	95% CI (LB)	95% CI (UB)
Location: Identity				
var(Constant)	0.307	0.177	0.099	0.95
var(Residual)	2.085	0.098	1.901	2.286

LR test vs. linear model: chibar2(01) = 55.70; Prob $\geq$ chibar2 = 0.00

8.6.2 Effect of Pedestrian Behavioural Characteristics

Pedestrian behavioural characteristics such as crossing in a group were observed to be a significant factor in risky crossing behaviour reported by past studies. The group's effect on the safety margin showed no significant impact in the case of one-way crosswalks. In contrast, for the two-way crosswalks, pedestrians in a group took more caution by taking a larger safety margin. It was observed that pedestrians took 11% larger safety margin when crossing in a group (8.2) compared to crossing alone. One possible reason could be that pedestrians in a group were more concerned about the group members' safety, thus deterring themselves from taking unnecessary risks.

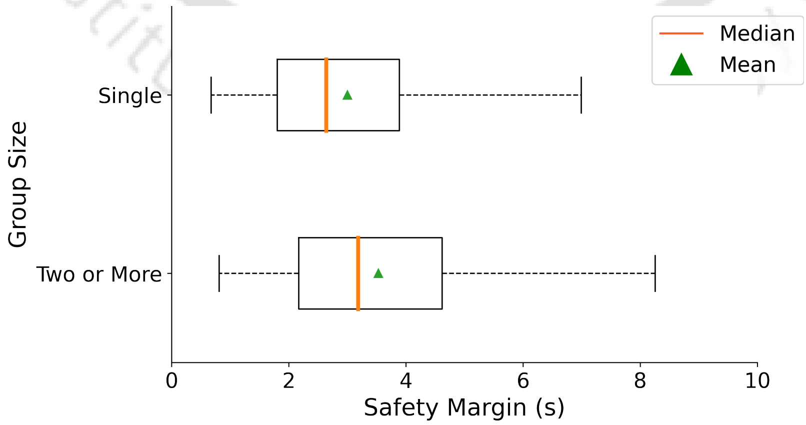


FIGURE 8.2: Impact of group/platoon size on Pedestrians Safety Margin (two-way crosswalks)

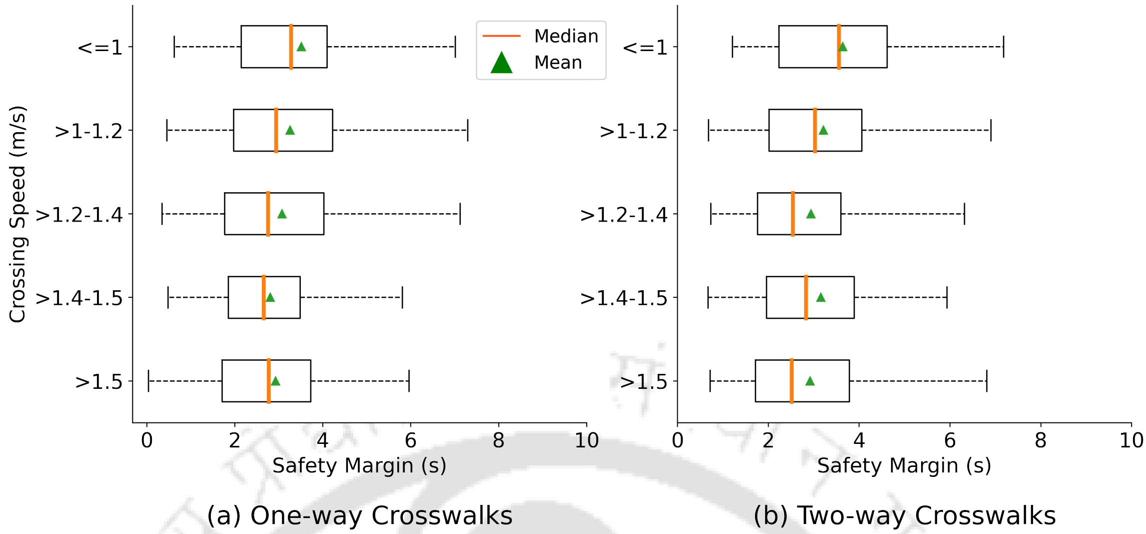


FIGURE 8.3: Impact of crossing speed on Pedestrians Safety Margin.

Further, the crossing speed variable revealed that pedestrians who can increase their speed during crossing could adopt a smaller safety margin. This behaviour is observed especially among young pedestrians who can increase their speed more frequently to adopt small vehicular gaps (Figure 8.3). It was observed that in one-way and two-way crosswalks, pedestrians with a crossing speed of >1.5 m/s accepted 30.6% and 29.5% smaller safety margins compared to reference crossing speed ($\leq$1 m/s).

8.6.3 Effect of Exposure

Pedestrian exposure to traffic is investigated using three variables when a pedestrian encounters an oncoming vehicle. The variables are the vehicle's position, type of vehicle and direction of its arrival (applicable for two-way crosswalks only). For one-way crosswalks, it has been observed that pedestrians took a smaller safety margin at the end lane. In contrast, for two-way crosswalks, no significant influence of position was observed. The reason might be that the end lane is near the footpath rest area; pedestrians see this as a safer zone than the middle lane of the road. Thus, to attain the reward of reaching the so-called safe zone, pedestrians undertook more risks and accepted small gaps.

Further, results from both the models revealed that as the vehicle size increases, pedestrians' safety margin size requirement for safe crossing also increases, which can also be validated from Figure 8.4. Pedestrians took 25.5% (one-way crosswalks) and 18.9% (two-way crosswalks) largest safety margin when the oncoming vehicle was a bus compared to a two-wheeler (base-case). This result is similar to the past study conducted by Avinash et al. (2018) and Zhuang and Wu (2011). In addition, the results on interaction with turning vehicles illustrate that pedestrians took the largest safety margin with right-turning vehicles (14.4% larger). The reason could be that right-turning vehicles' trajectory seemed almost a blind spot for pedestrians while crossing the road. Thus, interaction with the right-turning vehicle could have felt like a surprise; thus, to be on the safe side, pedestrians took more caution.

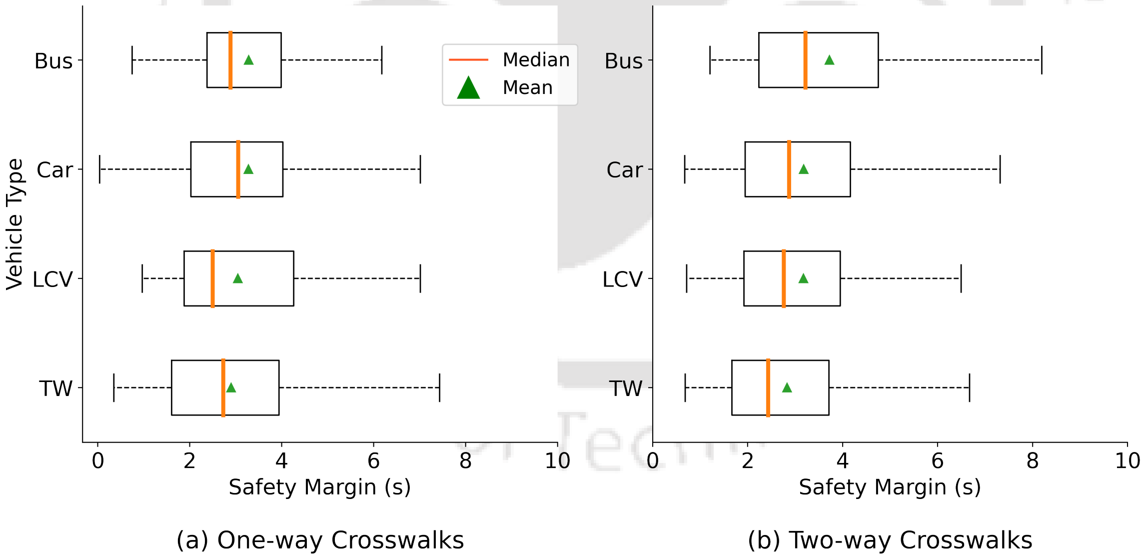


FIGURE 8.4: Impact of vehicle type on Pedestrian Safety Margin.

8.6.4 Effect of State of Crossing

The distraction variable was included in the two-way PSM model based on the Kruskal-Wallis significant test statistics. The distraction variable's coefficient revealed that only the "headphones use" category has a significant influence on PSM (refer Table 8.4 and Figure 8.5). The estimate revealed that compared to pedestrians without any distraction (base category), those engaged in music listening via headphones took a 25.1% smaller safety margin. The reason behind such behaviour could be that headphones capture pedestrians' attention while obstructing traffic noise. This reduces the existing traffic awareness. Further, it makes pedestrians envision being in a natural environment, withdrawing their alertness from the situation (Walker et al., 2012). Koh and Wong (2014) studied signal violation behaviour in Singapore and observed no significant influence of distraction on gap acceptance behaviour.

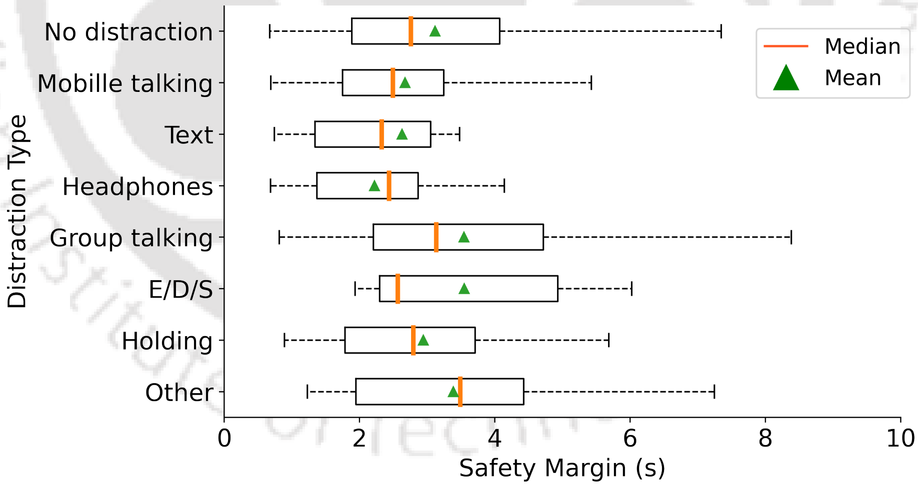


FIGURE 8.5: Impact of distraction on Pedestrian Safety Margin (Two-way crosswalks).

8.6.5 Effect of Cautionary/Glance Behaviour

Pedestrian cautionary or glance behaviour is an essential factor that provides important information to pedestrians, thus helping them plan a better and safer crossing decision. In both, the models, "before crossing glance", showed no significant influence on PSM. The reason could be that a "before crossing traffic glance" might have been useful only when pedestrians encountered the traffic in the initial lane (first lane). The difficulty might have escalated if they encountered the gap at mid-lane and end-lane; after looking at the vehicle by the time pedestrians reached at mid/end lane, the vehicle's dynamic could have changed, turning into an unsafe crossing. This can also be validated by the "during crossing glance" variable, which significantly impacted PSM. The results revealed that pedestrians accepted a 13.3% and 12.9% smaller safety margin when performing a "during crossing glance" at one-way and two-way crosswalks, respectively. The "during crossing glance" helped pedestrians keep track of the vehicular dynamics while rolling over one lane to another, thus, adopting a smaller safety margin (refer Figure 8.6).

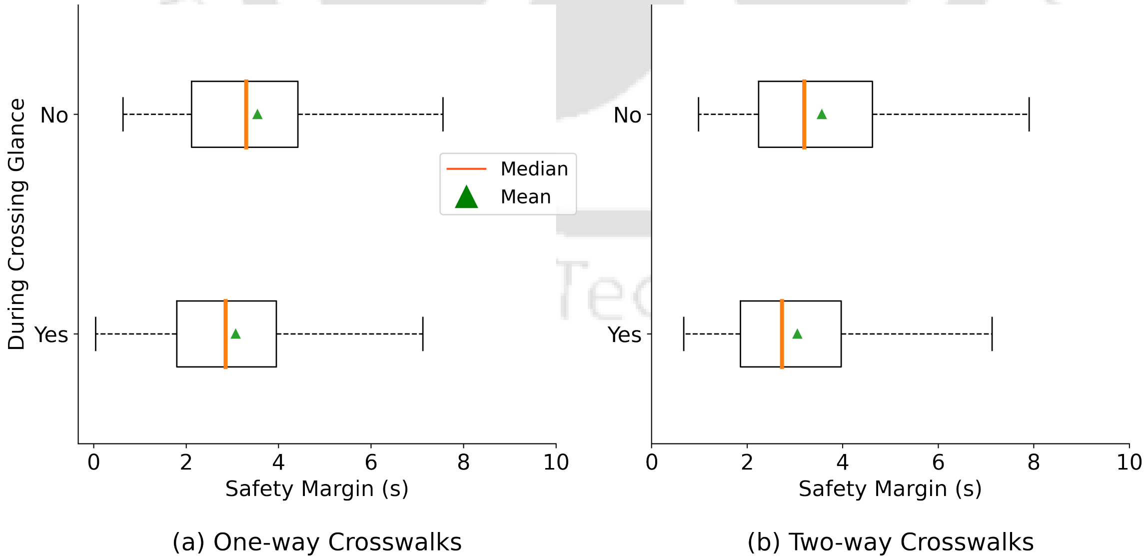


FIGURE 8.6: Impact of during crossing traffic glance on Pedestrian Safety Margin.

8.6.6 Effect of Traffic Characteristics

The number of traffic plying per cycle is also found to be negatively associated with the safety margin (refer Table 8.3 and Table 8.4), which is elucidated by the premise that when the number of vehicles increases, the safe gap between vehicles decreases. During high traffic volume cycles, pedestrians willing to take crossing risks accept a small safety margin to complete the manoeuvre, increasing their chances of crash risk.

8.7 Practical Implications from the PSM Results

- The impact of pedestrian behavioural characteristics on pedestrian safety has been identified, and these results may be useful to transportation engineers in similar kinds of roadway conditions to identify unsafe pedestrian intersection crosswalks.

- The impact of vehicle type (large vehicles) identified in the current study may be useful to transport planners. Intersections where traffic flow is comprised of a large number of heavy vehicles, such as buses and trucks, require special attention.

- The current study showed that "during crossing glance" is one of the significant parameters that guides pedestrians during the road crossing process. Thus, educating the young population about before crossing traffic glance (looking left and right) and during crossing traffic glance might help them cross the road safely.

- Pedestrians' movement needs to be separated entirely by providing foot-over bridges or underpasses at intersections with intense traffic to eliminate pedestrian-vehicle interactions.

- The study results may also be useful for developing new design standards for new pedestrian facilities where similar kind of roadways and traffic prevails.

8.8 Summary

The pedestrian safety margin (PSM) model presented in this chapter aims to identify the significant factors that influence PSM. The variables such as 'gender', 'age', 'traffic look before crossing', position where safety margin has been taken (one-way model), and different 'distractors' (one-way model) were found to be insignificant. The results show that pedestrian behavioural characteristics such as 'group size' (two-way model) and 'crossing speed' affect the safety margin value maintained by the pedestrians. Interaction with large vehicles and the presence of right-turning traffic increased safety margin, while "during crossing traffic glance" and through traffic per cycle were observed to reduce the pedestrian safety margin values.

The conclusions and research contributions of the current work are discussed in the next chapter.

Chapter 9

Conclusions and Practical Implications

9.1 Summary

In the current study, the methodologies were proposed to develop signal violation, distracted crossing, waiting behaviour and safety margin models. A video graphic and a questionnaire survey (for distraction) were conducted at selected signalised intersection crosswalks in Kolkata, India. The required quantitative data (such as pedestrian behavioural and traffic characteristics) were extracted from collected video footage. Initially, the influence of social and non-social factors on signal violation behaviour was studied at three one-way crosswalks and relationships were established. Later, the influence of various distractors (such as mobile phone talking, texting, headphones use, group talking and eating/drinking/smoking) was studied to understand whether these distractors promote unsafe road crossing behaviour. Afterwards, pedestrian waiting behaviour and the influence of covariates' on signal violation were studied using time-to-event analysis at two-way crosswalks. Finally, pedestrian and vehicle interaction was quantified using the safety margin approach. The significant findings of the research are classified into four categories: i) pedestrian signal violation behaviour model, ii) pedestrian distracted road crossing behaviour model, iii) pedestrian waiting behaviour survival model, and iv) pedestrian safety margin model at signalised intersection crosswalks and summarised in the following sections.

9.2 Study Findings

This study addressed the pedestrian unsafe road crossing behaviour at signalised intersection crosswalks under mixed traffic conditions in Kolkata, India. Major conclusions drawn from this study are presented below.

9.2.1 Pedestrian Signal Violation Behaviour Model

The specific conclusions drawn from the data collection and analysis of social and non-social factors influencing signal violation behaviour are as follows:

(i) The study provided empirical evidence that social and non-social factors influenced pedestrians' red-light violation behaviour. This is a significant finding because both factors can be used to understand signal violation behaviour. The red-light violation could be utilised as a surrogate measure to identify unsafe intersections in Kolkata city.

(ii) Regarding the group effect, the tendency to wait on the red-light phase was greater when more people were waiting at the curb, either when a pedestrian arrived, or other pedestrians joined after the arrival of the subject pedestrian. The present finding indicates a group's power to influence its members to obey the law positively. Media campaigns should utilise this power for educational purposes by emphasising the positive value of social control and its benefits.

(iii) The chances of pedestrians' signal violation are significantly high if a pedestrian finds that a significant number of neighbours successfully crossed the road, violating the red-signal phase.

(iv) The number of neighbours crossing the street in the red-light phase is a significant influencer for signal violation, compared to binary selection (someone crossed in the red-light phase or did not cross) used in past studies.

(v) The model outcome also revealed that compared to 0-20 seconds waiting time (reference category), if pedestrians experienced a longer waiting time of 20-40 seconds or >40 seconds for safe crossing, they became impatient, and their illegal crossing probability was increased by 12.8% and 28.4% respectively. The present finding points to the deficiencies in signal design and infrastructure planning.

9.2.2 Pedestrian Distracted Road Crossing Behaviour Model

In the era of digital devices, mobile phone use has proliferated in developing countries and caused safety compromises during road crossings. As vehicle drivers and pedestrians use a shared space for movement (especially at intersections); thus, similar to driver distraction, the pedestrian could also engage in such distraction, which reduces crossing safety. Thus, to understand the behavioural difference between pedestrians with and without distraction in this study, a detailed investigation has been conducted, considering three one-way crosswalks in Kolkata, India. The following conclusions are drawn from different pedestrian distraction model results.

(i) Questionnaire survey response showed that among all reported mobile use, the primary segment of the respondents was reported to be distracted from frequent mobile phone talking (81.2%) and when it was work-related (41%).

(ii) About 13.7% (61) of respondents was subjected to at least one near-miss in the past, and 4.5% (20) experienced at least one accident.

(iii) Pedestrians of the young adult age group (18-29 years) were more likely to talk on their mobile phones, while the teenage group (<18 years) was more involved in texting than any other age group when crossing the road at intersection crosswalks.

(iv) Pedestrians, with a wait time of one or more seconds, were 6% more likely to talk on a mobile phone compared to others during the crossing process.

(v) The pedestrian distracted by talking on the mobile phone, texting, eating/-drinking/smoking, and group talking was less likely to perform 'glance/cautionary behaviour' before and during the crossing.

(vi) Pedestrians engaged in texting were 7.9% more likely to cross in the red-signal phase compared to the green walk-signal phase.

(vii) In comparison to pedestrians without distraction, mobile-talking pedestrians were 4.5% more likely to 'nearly hit/bump' into another oncoming pedestrian.

(viii) Pedestrians crossing alone were found to perform more mobile talking and texting due to less social interactions and boredom as compared to pedestrians in groups of two or more.

(ix) Many variables in the "holding a phone in hand" analysis did not reach significance, indicating that pedestrians holding a phone (not using it) did not contribute to distraction. Thus, it should not be treated as any sort of distraction.

(x) Pedestrians involved in texting or eating/drinking/smoking walk considerably slower than their undistracted counterparts, which might increase the traffic exposure time and subsequently affect their safety.

9.2.3 Pedestrian Waiting Behaviour Survival Model

Pedestrian waiting time has a significant influence on risk-taking behaviour at signalised intersection crosswalks. The waiting duration of pedestrian crossing through the intersection crosswalks was modelled using a hazard-based duration approach. The conclusions drawn from the time-to-event analysis results are as follows:

(i) The study result revealed that the hazard-based duration approach could be a very useful technique for understanding the signal violation behaviour over a given duration of the waiting period.

(ii) Log-Normal distribution best represents the waiting duration of pedestrians at intersection crosswalks in Kolkata.

(iii) The high danger rate is shown by the short waiting time (less than 3 seconds), where 49.5% of pedestrians crossed the road immediately after red-phase arrival, indicating a pre-determined crossing decision or familiarity with the intersection.

(iv) The pedestrian endurance towards waiting was reflected in the span of 36.7 seconds, over which 75% of the pedestrians violated the signal.

(v) The pedestrians carrying luggage were observed to be more likely to violate traffic signal rules (survival time dropped by 38%). Hence, intersections with a high volume of such pedestrians (like intersections near public transportation terminals) require additional attention.

(vi) Group talking social distraction reduces signal violation tendency and increases waiting for endurance for the green light.

(vii) Pedestrian glance is an important predictor of signal violation. The study results revealed that pedestrians who looked at traffic before initiating crossing were less likely to violate the signal (survival wait time increased by 1131.72%).

(viii) The model outcome also suggested that a longer cycle length increases the probability of signal violation. This is an important finding, capturing broader planning, design and enforcement issues influencing pedestrian risk-taking behaviour.

9.2.4 Pedestrian Safety Margin Model

The pedestrian safety margin is a good quantitative dummy measure of unsafe crossing behaviour at signalised intersection crosswalks. In the present study, the pedestrian safety margin (PSM) at intersection crosswalks was modelled using the Multiple Linear Regression (for one-way crosswalks) and Random Intercept Linear model (for two-way crosswalks), and significant covariates were identified. The conclusions drawn from the PSM model are as follows:

(i) The result showed that pedestrian gender and age do not have any significant influence on pedestrian safety margin.

(ii) Pedestrians took 11% (two-way crosswalks) larger safety margin when crossing in a group compared to crossing alone.

(iii) Pedestrians took 25.5% (one-way crosswalks) and 18.9% (two-way crosswalks) larger safety margin when crossing against a bus compared to a two-wheeler.

(iv) The model result suggested that pedestrians with higher crossing speeds can accept a smaller safety margin.

(v) An exciting finding is that pedestrians took a smaller safety margin at the end lane to complete the crossing manoeuvre.

(vi) Model result deliberated that pedestrians took more caution in the presence of turning traffic than through traffic.

(vii) The current study results showed that "during crossing glance" helped pedestrians to keep track of the vehicular dynamics while rolling over from one lane to the other. This helped them adopt smaller safety margins.

(viii) Pedestrians with higher crossing speeds showed riskier crossing behaviour, i.e., took a smaller safety margin.

9.3 Research Contribution

This study advances the state of the art in understanding pedestrian risk-taking behaviour at signalised intersection crosswalks. The study enhances and develops models using real-world empirical data. Moreover, the current research effort provides the following contributions:

- The study of signal violation risk-taking behaviour demonstrated the influence of social factors along with non-social factors that encourages risk-taking behaviour at intersection crosswalks. The research highlighted the positive and negative implications of having social information disseminated from other pedestrians in road crossing decision-making.

- This study explored distracted road-crossing behaviour to understand the influence of distraction on road-crossing safety. This provides an overall framework for the safety evaluation of intersection crosswalks regarding digital and social distraction.

- Time to event analysis (survival analysis) was used to model pedestrian waiting behaviour at intersection crosswalks. The study provided a good overview of survival analysis in pedestrian safety research. The application part could be extended to other parts of transportation research.

- The study also explored a proactive evaluation mechanism of pedestrian safety using the safety margin concept at signalised intersection crosswalks under mixed traffic conditions. It developed a realistic model considering pedestrian crossing behavioural characteristics on pedestrian safety evaluation.

9.4 Policy Implications

Due to the increase in pedestrian crashes at intersection crosswalks, transport policymakers face myriad challenges in reconciling unsafe road crossing behaviour. Clear directives and formulation of new policies based on pedestrian crossing behaviour provide an opportunity to reinforce the pedestrian infrastructure. In particular, the impact of individual risk-taking behaviour at crosswalks led to many crashes and fatalities in developing countries, India in particular. Recognising that reconfiguring or formulating transport policies should be based on users' current behaviour, several possible policy implications obtained from this research are summarised below.

Prioritising Pedestrians: In India, pedestrian infrastructures are lagging, therefore, pushing them to behave in an unsafe manner. For example, none of the selected sites in the current study had a well-designed refuse island, which forced pedestrians to wait at the median in an unsafe manner. Hence, building safe and updated pedestrian infrastructure should be one of the top priorities for policymakers that make pedestrians feel valued and safe. Thus, increasing trust and respect for the traffic law.

Forgiving Infrastructure: In the current study, the K-M estimate showed that 75% of pedestrians got impatient and crossed the road within 36.7 s. The finding pointed out the deficiencies in signal design and infrastructure planning. A similar outcome has been observed for the longer signal cycle lengths. One solution could be to adjust the red-light phase duration at busy intersections or provide grade-separated crossing facilities where the red-light phase length is too long. Past studies suggested keeping the red-phase length below 40 s to enhance pedestrian signal compliance (BAASS, 1989; Wang et al., 2011). In developing countries where traffic rules are lenient, this type of unsafe crossing is more observed within a short waiting duration. Thus, more studies must be conducted to develop a reasonable red-light phase length.

Moreover, in the current study, 49.5% observations waited only for a shorter duration (0-3 s), which could be due to the reason that the decision was predetermined, or the pedestrians were familiar with the intersection and thus decided not to wait further. It indicates even with the best possible measures put in place to stop unsafe crossing; there will still be pedestrians who would take the risk and cross the road in an unsafe manner. In the real-world scenario, exceptions exist; it would be unrealistic to expect every road user to abstain from unsafe behaviour. Thus, policymakers must make future infrastructures forgiving in nature so that the cost of unintended mistakes can be avoided. One example could be narrowing down the crosswalk section of the road, which would reduce the crossing duration as well as the pedestrian exposure time if entering the road during the red-signal phase, and also force drivers to reduce their speed near the intersection, which helps in reducing fatal crashes (Distefano and Leonardi, 2017). A before and after implementation has been illustrated in Figure 9.1. Further, flashing LED lights and Variable Message Signs (VMS) could be installed to reduce serious vehicle-pedestrian conflicts and improve drivers' yielding behaviour at intersection crosswalks (Hussain et al., 2021). In addition, traditional fixed time signals should be replaced with count-down signals at intersection crosswalks, as it reduces crashes involving pedestrians and vehicles (Pulugurtha et al., 2010).

Road Safety Education: Although improvements made in the design of a built environment to improve pedestrians' convenience are necessary, re-educating and motivating millions of pedestrians to make a long-lasting behavioural change should also be the goal of policymakers. In the current study, we observed that glance is one of the most important parameters that help pedestrians in crossing decision-making. Past studies on pedestrians' distracted road-crossing behaviour also highlighted the importance of glances in road-crossing behaviour (Bungum et al., 2005; Pešić et al., 2016). Thus, educating the young population about traffic glances (looking left and right) and taking precautions before making the crossing decision might help them cross safely.

Innovation in Road Safety Interventions: Policymakers and planners could

FIGURE 9.1: Human-oriented street design

take advantage of modern innovative solutions to develop innovative road safety interventions. For example, computer vision-enabled AI systems could be utilised to track red-light violations and to warn road users (Keller and Gavrila, 2014). Similarly, Agent-based frameworks could be utilised to evaluate pedestrian road crossing safety at intersection crosswalks (Zhu et al., 2021). Although the impact of distraction in the current study is not significant in the observed intersections, still wherever this type of behaviour is found to be unsafe during road crossing on those intersections, innovative interventions could be applied. For example, stencil reminder (refer Figure 9.2 (A)) could be installed at crosswalk waiting areas (Barin et al., 2018). In-ground LED lights, or audio-based systems in pathways, could be used to seek distracted pedestrians' attention (refer Figure 9.2 (B)) and warn them about the signal status (Larue et al., 2019). Further, public campaigns and advertisements (refer Figure 9.2 (C) and 9.2 (D)) could be utilised to warn pedestrians about the crossing risk while distracted (Mwakalonge et al., 2015).

(A) Stencil reminder installed at crosswalks in California, USA (Barin et al., 2018)

(B) In-ground LEDs installed at crosswalks in South Korea (Larue et al., 2019)

(C) The "E-Lane" in Philadelphia (Mwakalonge et al., 2015)

(D) Distracted walking street campaign ad in San Francisco (Mwakalonge et al., 2015)

FIGURE 9.2: Interventions to reduce distracted road crossing behaviour

Further, installing pedestrian push buttons and Accessible Pedestrian Signals (APS) in place of existing signals could help pedestrians cross the road safely (refer Figure 9.3). Moreover, additional buttons need to be introduced in the existing Accessible Pedestrian Signals (APS) that would extend the normal crossing time (walk time) to accommodate the need of pedestrians with disabilities (walking disability) or impaired vision that could help them cross the road safely.

(A) Pedestrian push button

(B) Accessible Pedestrian Signals (APS)

FIGURE 9.3: Pedestrian crossing assistance devices

Strict Enforcement: Traffic laws were enforced to prevent and protect road users, mainly pedestrians, worldwide. These rules and regulation protect pedestrians and safeguards against rush drivers, also instruct drivers that they must slow down their vehicle when approaching a pedestrian crossing facility and it also prevents drivers from parking his/her vehicle near a traffic light/pedestrian crossing facility/crosswalk. Even though these rules are in place, road users in developing countries and especially drivers, are less likely to comply with rules and regulations due to lack of enforcement, thus resulting in numerous traffic violations leading to many pedestrian-vehicle crashes. In the current study, we observed that in many pedestrian green phases (red phase for vehicles), vehicles occupied the pedestrian right of way (stopped over zebra crossing) while waiting for their green light, which forced pedestrians to move out of the zebra crossing, which is very common in Indian scenario. Therefore, in addition to building pedestrian-friendly infrastructures, it is also essential that law enforcement agencies strictly enforce traffic laws in the field to safeguard pedestrians and other road users. For example, at intersections, a speed table with raised crosswalks could be provided (refer Figure 9.4), which not only helps in reducing approaching vehicle speed during the vehicle green phase but also discourages drivers from occupying the pedestrian right of way (zebra crossing) due to raised nature of the crosswalk during vehicle red-phase (i.e., while waiting for vehicle green phase), (Distefano and Leonardi, 2017).

FIGURE 9.4: Speed table with raised crosswalk

9.5 Study Limitations and Scope for Future Work

Like any other study, this study is also not without limitations due to a shortage of resources and field setting. The study limitations and future research scopes are presented below.

- The pedestrians' road crossing behaviour was observed only during non-peak hours (11 am to 2 pm) to avoid the influence of traffic police (forcibly stopping pedestrians from signal violations). Thus, future research should include peak and non-peak hours to study pedestrians' unsafe road crossing behaviour.

- In signal violation research, social and non-social factors were studied on one-way crosswalks. In future research, the same approach could be extended to two-way crosswalks also.

- Research involving the interaction between variables (demographics of leader and follower pedestrian violating signal) could help better understand the pedestrian signal violation risk-taking behaviour.

- The current study did not account for the driver's behaviour, such as yielding behaviour, speeding and rush driving into consideration. Future research involving these explanatory variables can provide better insight into pedestrian risk-taking behaviour.

- To get an initial understanding of pedestrian distracted road crossing behaviour in developing countries like India, similar to past studies, the present study also considered crosswalks with one-step crossing (three one-way crosswalks). The current study methodology can be extended to two-way divided streets in terms of two-step crossing to understand more complex crossing behaviour, where pedestrians could use mobile phones in the first step of crossing (start to median) and might not use them on the second step (median to end) or vice-versa.

- Future research should also investigate how pedestrians utilise their hearing senses to perceive their environments (especially pedestrians who are holding their phones or texting and not wearing headphones). This would provide a better understanding of how situational awareness is reduced by mobile phone use or if pedestrians are relying on senses other than sight to inform them of traffic conditions.

- The current distraction study only measured the influence of various distractors on road crossing behaviour. In future research, distracted unsafe road crossing behavioural estimates could be incorporated with the "Level of Service" measures to identify unsafe signalised intersection crosswalks.

- Combining objective observational data (on all stages of pedestrian road crossing) with subjective data (questionnaire survey) from the same observed pedestrian would provide a better understanding of human factors involved in crossing behaviour at signalised intersection crosswalks.

- The observations under the current study were made only based on one city-level data, which may restrict the scope of the conclusions drawn from the current study. Thus, future studies need to extend the methodology to a wider variety of intersections and samples with a more diverse category of signalised and unsignalised intersections over different Indian provinces (commercial, residential, and educational) to make findings more generalised.

Although there are several possibilities for future research in this domain, the results obtained from this study can be of great help to researchers, policymakers, design practitioners, and engineers. These findings will also be useful for designing better interventions and reducing unsafe crossing behaviour.

Appendix A

Questionnaire Survey Format

Pedestrians' Mobile Phone Distraction and Unsafe Road Crossing Behaviour
User Perception Survey

Survey Permission Memo No. TP/ PA/1/55329 (Traffic Department Kolkata, 20.08.2018)

Surveyor's Name: ____________ Location: __________________ Date & Response Time: ____________

A. Demographic Characteristics

A1) Gender: 1) Male ☐ 2) Female ☐ **A2) Age:** 1) <18 ☐ 2) 18-29 ☐ 3) 30-45 ☐ 4) 46-60 ☐ 5) >60 ☐

A3) Occupation: 1) Govt. Employee ☐ 2) Private Job ☐ 3) Self Employed ☐ 4) Student ☐ 5) Others

B. Distracted Walking

B1) What **type of phone** do you regularly use?
1) Feature Phone ☐ 2) Smart phone without internet ☐ 3) Smartphone with internet ☐

B2) While **walking on/ crossing road** for **what purposes** do you **frequently use your mobile phone**? [**Multiple Choice Ques.**]

1) Calling	☐	6) Mail	☐
2) Typing	☐	7) Navigation	☐
3) Music	☐	8) Browsing	☐
4) Social	☐	9) Gaming	☐
5) PhotoVideo	☐	10) Others	☐

B3) In **what situations do you generally use mobile phone** while walking on/ crossing roads?

1) During non-stressful traffic condition	☐	3) No police in sight	☐	5) Personal (family related)	☐
2) Boredom	☐	4) If it is work related	☐	6) Others	☐

B4) Have you **ever experienced** an **accident** or **near-miss** (accident) **due to mobile phone use** while walking on/ crossing roads? **[If respondent's answer is *"yes"*, then whether it is an accident, near-miss or both]**
1) Accident ☐ 2) Near-miss ☐ 3) None ☐

Thank you for your participation

Figure A.1: Mobile phone distraction questionnaire survey format

Appendix B

Data Extraction Sheet

location	gender	pace	crossing_type	speed	wait_till_green	people_waiting	people_joining	cross_from_same	cycle_time_cat	traffic	sig_dep
1	0	1	0	4	1	2	15	1	0	26	0
1	0	1	0	4	0	11	17	2	0	26	0
1	1	0	0	3	0	12	11	2	0	26	0
1	0	1	0	4	0	13	1	2	0	26	1
1	0	0	0	2	1	1	12	1	2	40	1
1	1	0	0	4	1	9	2	1	2	40	1
1	1	0	0	1	1	8	10	1	2	40	1
1	0	0	0	2	1	1	14	9	2	40	1
1	0	0	1	1	0	2	1	2	2	40	1
1	0	0	0	2	0	2	1	9	2	40	1
1	0	0	1	1	0	1	3	2	1	45	1
1	0	1	1	4	1	10	11	1	1	45	1

FIGURE B.1: Data extraction sheet [includes raw data of red-light arriving pedestrians]

Where,

location: Survey Sites.

gender: Gender (Male/Female).

pace: Walking pace (Normal/Hurried).

crossing_type: Crossing path (Perpendicular/Oblique).

Speed: Crossing speed in m/s.

wait_till_green: Waiting duration (s) till green phase initiation.

people_waiting: Number of pedestrian waiting on arrival.

people_joining: Number of people joined after the candidate pedestrian's arrival.

cross_from_same: Number of pedestrian crossing from same direction during red-light phase.

cycle_time_cat: Cycle time category (0-100 s, >100-150 s and >150 s).

traffic: Number of traffic plying on the road per cycle.

sig_dep: Signal status on departure (green walk/ red do not walk).

References

Afshari, A., Ayati, E., Barakchi, M., 2021. Evaluating the effects of external factors on pedestrian violations at signalized intersections (a case study of Mashhad, Iran). IATSS Research 45, 234–240. doi:10.1016/j.iatssr.2020.10.004.

Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control 19, 716–723. doi:10.1109/TAC.1974.1100705.

Alexandre, M., Ruggiero, L., 2015. Illegal pedestrian crossing at signalised junctions in urban areas : The impact of spatial factors. Transportation Research Board, 94th Annual Meeting , 1–14doi:10.13140/2.1.1534.7520.

Alhajyaseen, W.K., Iryo-Asano, M., 2017. Studying critical pedestrian behavioral changes for the safety assessment at signalized crosswalks. Safety Science 91, 351–360. doi:10.1016/j.ssci.2016.09.002.

Avinash, C., Gore, N., Shriniwas, A., Gaurang, J., Manoranjan, P., 2020. Choice crossing behaviour model for Safety Margin of pedestrian at mid-blocks in India. Transportation Research Procedia 48, 2329–2342. doi:10.1016/j.trpro.2020.08.285.

Avinash, C., Jiten, S., Shriniwas, A., Gaurang, J., Manoranjan, P., 2018. Evaluation of pedestrian safety margin at mid-block crosswalks in India. Safety Science 119, 188–198. doi:10.1016/j.ssci.2018.12.009.

BAASS, K.G., 1989. Review of European and North American practice of pedestrian signal timing., in: Prepared for RTAC Annual Conference Calgary, Alberta. pp. 17–21.

Banducci, S.E., Ward, N., Gaspar, J.G., Schab, K.R., Crowell, J.A., Kaczmarski, H., Kramer, A.F., 2016. The Effects of Cell Phone and Text Message Conversations on Simulated Street Crossing. Human Factors: The Journal of the Human Factors and Ergonomics Society 58, 150–162. doi:10.1177/0018720815609501.

Bandura, A., 1977. Social learning theory.

Bandura, A., 1986. Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bansal, A., Goyal, T., Sharma, U., 2018. Pedestrian Safety on Crosswalks in India - Need of the Hour. Journal on Today's Ideas - Tomorrow's Technologies 6, 35–46. URL: https://jotitt.chitkara.edu.in/index.php/jotitt/article/view/19, doi:10.15415/jotitt.2018.61004.

Barin, E.N., McLaughlin, C.M., Farag, M.W., Jensen, A.R., Upperman, J.S., Arbogast, H., 2018. Heads Up, Phones Down: A Pedestrian Safety Intervention on Distracted Crosswalk Behavior. Journal of Community Health 0, 1–6. doi:10.1007/s10900-018-0488-y.

Baswail, A., Allinson, L., Goddard, P., Pfe, K., 2019. Adolescents ' Mobile Phone Use While Crossing the Road. Safety 5(2). doi:doi:10.3390/safety5020027.

Bentzen, B.L., Barlow, J.M., Bond, T., 2004. Challenges of unfamiliar signalized intersections for pedestrians who are blind: Research on safety. Transportation Research Record , 51–57doi:10.3141/1878-07.

Biswas, S., Ghosh, I., Chandra, S., 2017. Effect of Traffic Signal Countdown Timers on Pedestrian Crossings at Signalized Intersection. Transportation in Developing Economies 3, 2. URL: http://link.springer.com/10.1007/s40890-016-0032-7, doi:10.1007/s40890-016-0032-7.

Boroujerdian, A.M., Nemati, M., 2016. Pedestrian Gap Acceptance Logit Model in Unsignalized Crosswalks Conflict Zone. International Journal of Transportation Engineering 4, 87–96. doi:https://dx.doi.org/10.22119/ijte.2016.40534.

Brosseau, M., Zangenehpour, S., Saunier, N., Miranda-Moreno, L., 2013. The impact of waiting time and other factors on dangerous pedestrian crossings and violations at signalized intersections: A case study in Montreal. Transportation Research Part F: Traffic Psychology and Behaviour 21, 159–172. doi:10.1016/j.trf.2013.09.010.

Bungum, T.J., Day, C., Henry, L.J., 2005. The association of distraction and caution displayed by pedestrians at a lighted crosswalk. Journal of Community Health 30, 269–279. doi:10.1007/s10900-005-3705-4.

Cameron, D.P., Kalderstam, J., Zivich, P., Kuhn, B., Williamson, M., Fiore-Gartland, A., Moneda, L., WIlson, D., Parij, A., Stark, K., Anton, S., 2019. CamDavidsonPilon/lifelines: v0. 23.9. Zenodo.

Cao, Y., Ni, Y., Li, K., 2017. Effects of Refuge Island Settings on Pedestrian Safety Perception and Signal Violation at Signalized Intersections. TRB 96th Annual Meeting 750, 1–16.

Chaudhari, A., Arkatkar, S., Joshi, G., Parida, M., 2020. Exploring stage-wise pedestrian-crossing behavioral patterns at vulnerable urban midblocks: a perspective under heterogeneous traffic conditions. Journal of Transportation Safety and Security 12, 863–891. doi:10.1080/19439962.2018.1543375.

Chen, P.L., Pai, C.W., 2018. Smartphone gaming is associated with pedestrians' head-turning performances: An observational study of street-crossing behaviors at uncontrolled intersection in Taipei. International Journal of Sustainable Transportation 12, 12–18. doi:10.1080/15568318.2017.1321706.

Chen, S., Xing, J., Cao, Y., 2017. The Impact of Waiting Time on Pedestrian Violations at Signalized Intersections. Civil Engineering and Urban Planning: An International Journal (CiVEJ) 4, 01–13. doi:10.5121/civej.2017.4201.

Cherry, C., 1978. On human communication: a review, a survey, and a criticism. 3rd ed., London MIT Press., Troy, New York.

Chu, X., Baltes, M., 2001. Pedestrian mid-block crossing difficulty (No. NCTR-392-09). Technical Report. University of South Florida.

Ciaburro, G., 2018. Regression Analysis with R. Packt Publishing.

Commissioner, C., 2011. Census of India 2011. Technical Report. Ministry of Home Affairs.

Cooper, J.F., Schneider, R.J., Ryan, S., Co, S., 2012. Documenting Targeted Behaviors Associated with Pedestrian Safety. Transportation Research Record: Journal of the Transportation Research Board 2299, 1–10. doi:10.3141/2299-01.

Cox, D.R., 1972. Regression Models and Life-Tables. Journal of the Royal Statistical Society: Series B (Methodological) 34, 187–202. doi:10.1111/j.2517-6161.1972.tb00899.x.

Dandona, R., Kumar, G.A., Gururaj, G., James, S., Chakma, J.K., Thakur, J.S., Srivastava, A., Kumaresh, G., Glenn, S.D., Gupta, G., Krishnankutty, R.P., Malhotra, R., Mountjoy-Venning, W.C., Mutreja, P., Pandey, A., Shukla, D.K., Varghese, C.M., Yadav, G., Reddy, K.S., Swaminathan, S., Bekedam, H.J., Vos, T., Naghavi, M., Murray, C.J., Dhaliwal, R.S., Dandona, L., 2020. Mortality due to road injuries in the states of India: the Global Burden of Disease Study 1990–2017. The Lancet Public Health 5, e86–e98. doi:10.1016/S2468-2667(19)30246-4.

Das, D.K., 2019. 62 pedestrians die daily in India, up 84% in 4 years — India News - Times of India. URL: https://timesofindia.indiatimes.com/india/62-pedestrians-die-daily-in-india-up-84-in-4-years/articleshow/72101003.cms.

Dass, S., Singhal, D., Aggarwal, P., Kurukshetra, N.I.T., 2015. Study of Pedestrian Flow / Behavior on Indian Roads. IOSR Journal of Mechanical and Civil Engineering , 38–42.

Demiroz, Y.I., Onelcin, P., Alver, Y., 2015. Illegal road crossing behavior of pedestrians at overpass locations: Factors affecting gap acceptance, crossing times and overpass use. Accident Analysis and Prevention 80, 220–228. URL: http://dx.doi.org/10.1016/j.aap.2015.04.018, doi:10.1016/j.aap.2015.04.018.

Dhoke, A., Kumar, A., Ghosh, I., 2021. Hazard-Based Duration Approach to Pedestrian Crossing Behavior at Signalized Intersections. Transportation Research Record: Journal of the Transportation Research Board , 036119812110031doi:10.1177/03611981211003102.

Distefano, N., Leonardi, S., 2017. Effects of Speed Table, Chicane and Road Narrowing on Vehicle Speeds in Urban Areas. 6th International Symposium New Horizon 2017 of Transport and Communication .

Dommes, A., Granié, M.A., Cloutier, M.S., Coquelet, C., Huguenin-Richard, F., 2015. Red light violations by adult pedestrians and other safety-related behaviors at signalized crosswalks. Accident Analysis & Prevention 80, 67–75. doi:10.1016/j.aap.2015.04.002.

Faria, J.J., Krause, S., Krause, J., 2010. Collective behavior in road crossing pedestrians: the role of social information. Behavioral Ecology 21, 1236–1242. doi:10.1093/beheco/arq141.

Forsythe, M., Berger, W., 1973. Urban Pedestrian Accident Countermeasures Experimental Evaluation. volume 1. Appendix C: Behavioral Evaluation Summary Data, Biotechnology, Inc. Falls Church, VA; US Department of Transportation, Washington, DC.

Gillette, G., Fitzpatrick, K., Chrysler, S., Avelar, R., 2016. Effect of Distractions on a Pedestrian's Waiting Behavior at Traffic Signals: Observational Study. Transportation Research Record: Journal of the Transportation Research Board 2586, 111–119. doi:10.3141/2586-13.

Giraldeau, L.A., Valone, T.J., Templeton, J.J., 2002. Potential disadvantages of using socially acquired information. Philosophical Transactions of the Royal Society B: Biological Sciences 357, 1559–1566. doi:10.1098/rstb.2002.1065.

Gold, J.E., Rauscher, K.J., Zhu, M., 2015. A validity study of self-reported daily texting frequency, cell phone characteristics, and texting styles among young adults. BMC Research Notes 8, 1–7. doi:10.1186/s13104-015-1090-3.

Grambsch, P., Therneau, T., 1994. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81, 515–526. doi:10.1093/biomet/81.3.515.

Guo, H., Gao, Z., Yang, X., Jiang, X., 2011. Modeling pedestrian violation behavior at signalized crosswalks in China: A hazards-based duration approach. Traffic Injury Prevention 12, 96–103. doi:10.1080/15389588.2010.518652.

Guo, H., Wang, W., Guo, W., Jiang, X., Bubb, H., 2012. Reliability analysis of pedestrian safety crossing in urban traffic environment. Safety Science 50, 968–973. doi:10.1016/j.ssci.2011.12.027.

Haga, S., Sano, A., Sekine, Y., Sato, H., Yamaguchi, S., Masuda, K., 2015. Effects of using a Smart Phone on Pedestrians' Attention and Walking. Procedia Manufacturing 3, 2574–2580. doi:10.1016/j.promfg.2015.07.564.

Hamed, M.M., 2001. Analysis of pedestrians' behavior at pedestrian crossings. Safety Science 38, 63–82. doi:10.1016/S0925-7535(00)00058-8.

Hamidun, R., Ishak, S.Z., Endut, I.R., 2013. Assessing pedestrian crossing risk at signalised intersection. International Journal of Emerging Technology and Advanced Engineering 3, 31–35.

Hamidun, R., Kordi, N.E., Endut, I.R., Ishak, S.Z., Mohd Yusof, M.F., 2016. Pedestrian Potential Risk While Crossing at Signalized Intersections. Jurnal Teknologi 78, 15–21. doi:10.11113/jt.v78.8482.

Harrell, W.A., 1991. Factors Influencing Pedestrian Cautiousness in Crossing Streets. The Journal of Social Psychology 131, 367–372. doi:10.1080/00224545.1991.9713863.

Hatfield, J., Murphy, S., 2007. The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections. Accident Analysis & Prevention 39, 197–205. doi:10.1016/j.aap.2006.07.001.

Hirschi, T., 1969. Causes of Delinquency. Berkeley: University of California Press.

Holland, C., Hill, R., 2007. The effect of age, gender and driver status on pedestrians' intentions to cross the road in risky situations. Accident Analysis and Prevention 39, 224–237. doi:10.1016/j.aap.2006.07.003.

Holm, A., Jaani, J., Eensoo, D., Piksööt, J., 2018. Pedestrian behaviour of 6th grade Estonian students: Implications of social factors and accident-prevention education at school. Transportation Research Part F: Traffic Psychology and Behaviour 52, 112–119. doi:10.1016/j.trf.2017.11.005.

Horberry, T., Osborne, R., Young, K., 2019. Pedestrian smartphone distraction: Prevalence and potential severity. Transportation Research Part F: Traffic Psychology and Behaviour 60, 515–523. doi:10.1016/j.trf.2018.11.011.

Horrey, W.J., Wickens, C.D., 2006. Examining the impact of cell phone conversations on driving using meta-analytic techniques. Human Factors 48, 196–205. doi:10.1518/001872006776412135.

Hsiao, M., Malhotra, A., Thakur, J.S., Sheth, J.K., Nathens, A.B., Dhingra, N., Jha, P., 2013. Road traffic injury mortality and its mechanisms in India: Nationally representative mortality survey of 1.1 million homes. BMJ Open 3. doi:10.1136/bmjopen-2013-002621.

Hussain, Q., Alhajyaseen, W.K., Pirdavani, A., Brijs, K., Shaaban, K., Brijs, T., 2021. Do detection-based warning strategies improve vehicle yielding behavior at uncontrolled midblock crosswalks? Accident Analysis and Prevention 157, 106166. doi:10.1016/j.aap.2021.106166.

Hyman, I.E., Boss, S.M., Wise, B.M., McKenzie, K.E., Caggiano, J.M., 2009. Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. Applied Cognitive Psychology 24, 597–607. doi:10.1002/acp.1638.

IRC:103, 2012. Guidelines For Pedestrian Facilities. Technical Report. Indian Road Congress.

Iryo-Asano, M., Alhajyaseen, W.K., Nakamura, H., 2015. Analysis and Modeling of Pedestrian Crossing Behavior during the Pedestrian Flashing Green Interval. IEEE Transactions on Intelligent Transportation Systems 16, 958–969. doi:10.1109/TITS.2014.2346154.

Jain, A., Gupta, A., Rastogi, R., 2014. Pedestrian Crossing Behaviour Analysis at Intersections. International Journal for Traffic and Transport Engineering 4, 103–116.

Jensupakarn, A., Kanitpong, K., 2018. Influences of motorcycle rider and driver characteristics and road environment on red light running behavior at signalized intersections. Accident Analysis and Prevention 113, 317–324. URL: https://doi.org/10.1016/j.aap.2018.02.007, doi:10.1016/j.aap.2018.02.007.

Jiang, K., Ling, F., Feng, Z., Ma, C., Kumfer, W., Shao, C., Wang, K., 2018. Effects of mobile phone distraction on pedestrians' crossing behavior and visual attention allocation at a signalized intersection: An outdoor experimental study. Accident Analysis and Prevention 115, 170–177. doi:10.1016/j.aap.2018.03.019.

Kadali, B.R., Vedagiri, P., 2013. Modelling pedestrian road crossing behaviour under mixed traffic condition. European Transport - Trasporti Europei .

Kadali, B.R., Vedagiri, P., Rathi, N., 2015. Models for pedestrian gap acceptance behaviour analysis at unprotected mid-block crosswalks under mixed traffic conditions. Transportation Research Part F: Traffic Psychology and Behaviour 32, 114–126. doi:10.1016/j.trf.2015.05.006.

Kaplan, E.L., Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association 53, 457–481. doi:10.1080/01621459.1958.10501452.

Keller, C.G., Gavrila, D.M., 2014. Will the pedestrian cross? A study on pedestrian path prediction. IEEE Transactions on Intelligent Transportation Systems 15, 494–506. doi:10.1109/TITS.2013.2280766.

Koepsell, T., 2002. Crosswalk Markings and the Risk of Pedestrian–Motor Vehicle Collisions in Older Pedestrians. JAMA 288, 2136. doi:10.1001/jama.288.17.2136.

Koh, P., Wong, Y., 2014. Gap acceptance of violators at signalised pedestrian crossings. Accident Analysis & Prevention 62, 178–185. doi:10.1016/j.aap.2013.09.020.

Kumar, A., Paul, M., Ghosh, I., 2019. Analysis of Pedestrian Conflict with Right-Turning Vehicles at Signalized Intersections in India. Journal of Transportation Engineering, Part A: Systems 145, 1–12. doi:10.1061/jtepbs.0000239.

Kumar, S., Toshniwal, D., 2015. Analysing road accident data using association rule mining, in: 2015 International Conference on Computing, Communication and Security (ICCCS), IEEE. pp. 1–6. doi:10.1109/CCCS.2015.7374211.

Larue, G.S., Watling, C.N., Black, A.A., Wood, J.M., Khakzar, M., 2019. Pedestrians distracted by their smartphone: Are in-ground flashing lights catching their attention? A laboratory study. Accident Analysis and Prevention 134, 105346. doi:10.1016/j.aap.2019.105346.

Leeper, T.J., 2018a. Interpreting regression results using average marginal effects with R's margins. Reference manual. URL: https://cran.r-project.org/web/packages/margins/vignettes/TechnicalDetails.pdf.

Leeper, T.J., 2018b. margins: Marginal Effects for Model Objects. R package version 0.3.23.

Li, Y., Fernie, G., 2010. Pedestrian behavior and safety on a two-stage crossing with a center refuge island and the effect of winter weather on pedestrian compliance rate. Accident Analysis & Prevention 42, 1156–1163. doi:10.1016/j.aap.2010.01.004.

Lichenstein, R., Smith, D.C., Ambrose, J.L., Moody, L.A., 2012. Headphone use and pedestrian injury and death in the United States: 2004–2011. Injury Prevention 18, 287–290. doi:10.1136/injuryprev-2011-040161.

Lin, M.I.B., Huang, Y.P., 2017. The impact of walking while using a smartphone on pedestrians' awareness of roadside events. Accident Analysis & Prevention 101, 87–96. doi:10.1016/j.aap.2017.02.005.

Lin, P.S., Kourtellis, A., Zhang, Y., 2015. Understanding of Safety Impact and Potential Causes of Distracted Walking and Biking, in: Presentation, pp. 1–8.

Liu, Y.C., Tung, Y.C., 2014. Risk analysis of pedestrians' road-crossing decisions: Effects of age, time gap, time of day, and vehicle speed. Safety Science 63, 77–82. doi:10.1016/j.ssci.2013.11.002.

Lobjois, R., Cavallo, V., 2007. Age-related differences in street-crossing decisions: The effects of vehicle speed and time constraints on gap selection in an estimation task. Accident Analysis and Prevention 39, 934–943. doi:10.1016/j.aap.2006.12.013.

Marisamynathan, S., Vedagiri, P., 2018. Estimation of Pedestrian Safety Index Value at Signalized Intersections Under Mixed Traffic Conditions. Transportation in Developing Economies 4, 5. doi:10.1007/s40890-018-0058-0.

Ministry of Statistics and Programme Implementation, 2017. Youth in india 2017. Technical Report. Ministry of Statistics and Programme Implementation. URL: http://mospi.nic.in/sites/default/files/publication_reports/Youth_in_India-2017.pdf.

Mohan, D., Tiwari, G., Bhalla, K., 2020. Road safety in India status report. Technical Report May. Indian Institute of Technology Delhi, India. URL: http://tripp.iitd.ernet.in/assets/publication/Road_Safety_in_India_2019update1.pdf.

Mohd Syazwan, S., Baba, M.D., Nur Zarifah, H., Aqbal Hafeez, A., Noor Faradila, P., 2017. Prevalence of distracted pedestrians while crossing: a study of Malaysia's situation. MATEC Web of Conferences 90, 01031.

MoRTH, 2018. Road Accidents in India. Technical Report. Ministry of Road Transport & Highways (Transport Research Wing). New Delhi.

Moyano Diaz, E., 2002. Theory of planned behavior and pedestrians' intentions to violate traffic regulations. Transportation Research Part F: Traffic Psychology and Behaviour 5, 169–175. doi:10.1016/S1369-8478(02)00015-3.

Mukherjee, D., Mitra, S., 2017. What affects fatal pedestrian crashes at urban signalized intersections? A study on pedestrian behavior at signalized intersections in Kolkata using video-graphic survey., in: Road Safety and Simulation International Conference, Road Safety Simulation International Conference, Hague, Netherlands. p. 17–19.

Mukherjee, D., Mitra, S., 2019. Impact of Road Infrastructure Land Use and Traffic Operational Characteristics on Pedestrian Fatality Risk: A Case Study of Kolkata, India. Transportation in Developing Economies 5, 1–9. URL: https://doi.org/10.1007/s40890-019-0077-5, doi:10.1007/s40890-019-0077-5.

Mukherjee, D., Mitra, S., 2020. A comprehensive study on factors influencing pedestrian signal violation behaviour: Experience from Kolkata City, India. Safety Science 124, 104610. doi:10.1016/j.ssci.2020.104610.

Mwakalonge, J., Siuhi, S., White, J., 2015. Distracted walking: Examining the extent to pedestrian safety problems. Journal of Traffic and Transportation Engineering (English Edition) 2, 327–337. doi:https://doi.org/10.1016/j.jtte.2015.08.004.

Nasar, J.L., Troyer, D., 2013. Pedestrian injuries due to mobile phone use in public places. Accident Analysis & Prevention 57, 91–95. doi:10.1016/j.aap.2013.03.021.

Neider, M.B., McCarley, J.S., Crowell, J.A., Kaczmarski, H., Kramer, A.F., 2010. Pedestrians, vehicles, and cell phones. Accident Analysis & Prevention 42, 589–594. doi:10.1016/j.aap.2009.10.004.

Nukta Ramadani, H., Rahmani, H., Gazali, A., 2018. Study of efficiency pedestrian bridge crossing in the road of Pangerang Antasari, Banjarmasin. MATEC Web of Conferences 181, 06009. doi:10.1051/matecconf/201818106009.

Onelcin, P., Alver, Y., 2015. Illegal crossing behavior of pedestrians at signalized intersections: Factors affecting the gap acceptance. Transportation Research Part F: Traffic Psychology and Behaviour 31, 124–132. URL: http://dx.doi.org/10.1016/j.trf.2015.04.007.

Onelcin, P., Alver, Y., 2017. The crossing speed and safety margin of pedestrians at signalized intersections. Transportation Research Procedia 22, 3–12. URL: https://linkinghub.elsevier.com/retrieve/pii/S2352146517301369, doi:10.1016/j.trpro.2017.03.002.

Ortiz, N.C., Ramnarayan, M., Mizenko, K., 2017. Distraction and road user behavior: An observational pilot study across intersections in Washington, D.C. Journal of Transport and Health 7, 13–22. URL: https://doi.org/10.1016/j.jth.2017.05.362.

Osborne, R., Horberry, T., Young, K.L., 2020. Pedestrian distraction from Smartphones: An end-user perspective on current and future countermeasures. Transportation Research Part F: Traffic Psychology and Behaviour 73, 348–361. URL: https://doi.org/10.1016/j.trf.2020.07.007https://linkinghub.elsevier.com/retrieve/pii/S136984782030468X, doi:10.1016/j.trf.2020.07.007.

Oxley, J.A., Ihsen, E., Fildes, B.N., Charlton, J.L., Day, R.H., 2005. Crossing roads safely: An experimental study of age differences in gap selection by pedestrians. Accident Analysis & Prevention 37, 962–971. doi:10.1016/j.aap.2005.04.017.

Pawar, D.S., Patil, G.R., 2015. Pedestrian temporal and spatial gap acceptance at mid-block street crossing in developing world. Journal of Safety Research 52, 39–46. URL: http://dx.doi.org/10.1016/j.jsr.2014.12.006.

Pekker, A., Botos, A., Rusznyak, A., Koltai, J., Kurti, J., Kamaras, K., 2011. The Impact of Driver Inattention on Crash Risk: An Analysis using the 100-car Naturalistic Driving Study. Journal of Physical Chemistry Letters 2, 2079–2082. doi:10.1021/jz200899d.

Pele, M., Bellut, C., Debergue, E., Gauvin, C., Jeanneret, A., Leclere, T., Nicolas, L., Pontier, F., Zausa, D., Sueur, C., 2017. Cultural influence of social information use in pedestrian road-crossing behaviours. Royal Society Open Science 4, 160739. doi:10.1098/rsos.160739.

Pešić, D., Antić, B., Glavić, D., Milenković, M., 2016. The effects of mobile phone use on pedestrian crossing behaviour at unsignalized intersections – Models for predicting unsafe pedestrians behaviour. Safety Science 82, 1–8. doi:10.1016/j.ssci.2015.08.016.

Praveen Kumar N., C., N., Saranya, R., N. H., S., 2018. Impact of pedestrian crossing behavior at intersections on road traffic accidents in Shivamogga, Karnataka: an observational study. International Journal of Community Medicine And Public Health 5, 2437. doi:10.18203/2394-6040.ijcmph20182173.

Priyadarshini, P., Mitra, S., 2018. Investigating Pedestrian Risk Factors Leading to Pedestrian Fatalities in Kolkata City Roads. Transportation in Developing Economies 4, 1–11. URL: http://dx.doi.org/10.1007/s40890-017-0054-9, doi:10.1007/s40890-017-0054-9.

Pulugurtha, S.S., Desai, A., Pulugurtha, N.M., 2010. Are pedestrian countdown signals effective in reducing crashes? Traffic Injury Prevention 11, 632–641. doi:10.1080/15389588.2010.498845.

Pulugurtha, S.S., Srirangam, L.P., 2021. Pedestrian safety at intersections near light rail transit stations. Public Transport URL: https://doi.org/10.1007/s12469-021-00276-y, doi:10.1007/s12469-021-00276-y.

Ralph, K., Girardeau, I., 2020. Distracted by "distracted pedestrians"? Transportation Research Interdisciplinary Perspectives 5. doi:10.1016/j.trip.2020.100118.

Rankavat, S., Tiwari, G., 2016. Pedestrians risk perception of traffic crash and built environment features – Delhi, India. Safety Science 87, 1–7. doi:10.1016/j.ssci.2016.03.009.

Ren, G., Zhou, Z., Wang, W., Zhang, Y., Wang, W., 2011. Crossing Behaviors of Pedestrians at Signalized Intersections: Observational Study and Survey in China. Transportation Research Record: Journal of the Transportation Research Board 2264, 65–73. doi:10.3141/2264-08.

Retting, R.A., Chapline, J.F., Williams, A.F., 2002. Changes in crash risk following re-timing of traffic signal change intervals. Accident Analysis and Prevention 34, 215–220. doi:10.1016/S0001-4575(01)00016-1.

Rosenbloom, T., 2009. Crossing at a red light: Behaviour of individuals and groups. Transportation Research Part F: Traffic Psychology and Behaviour 12, 389–394. doi:10.1016/j.trf.2009.05.002.

Russo, B.J., James, E., Aguilar, C.Y., Smaglik, E.J., 2018. Pedestrian Behavior at Signalized Intersection Crosswalks: Observational Study of Factors Associated with Distracted Walking, Pedestrian Violations, and Walking Speed. Transpotation Research record doi:10.1177/0361198118759949.

Schober, M.F., Conrad, F.G., Fricker, S.S., 2004. Misunderstanding standardized language in research interviews. Applied Cognitive Psychology 18, 169–188. doi:10.1002/acp.955.

Schoenfeld, D., 1982. Partial residuals for the proportional hazards regression model. Biometrika 69, 239–241. doi:10.1093/biomet/69.1.239.

Schwarz, G., 1978. Estimating the Dimension of a Model." The Annals of Statistics. Statistics 6, 461–464. URL: http://www.jstor.org/stable/2958889.

Schwebel, D.C., Stavrinos, D., Byington, K.W., Davis, T., O'Neal, E.E., de Jong, D., 2012. Distraction and pedestrian safety: How talking on the phone, texting, and listening to music impact crossing the street. Accident Analysis &

Prevention 45, 266–271. URL: https://linkinghub.elsevier.com/retrieve/pii/S0001457511001965, doi:10.1016/j.aap.2011.07.011.

Shaaban, K., Muley, D., Mohammed, A., 2019. Modeling pedestrian gap acceptance behavior at a six-lane urban road. Journal of Transportation Safety and Security 13, 842–859. URL: https://doi.org/10.1080/19439962.2019.1691100, doi:10.1080/19439962.2019.1691100.

Simmons, S.M., Caird, J.K., Ta, A., Sterzer, F., Hagel, B.E., 2020. Plight of the distracted pedestrian: a research synthesis and meta-analysis of mobile phone use on crossing behaviour. Injury Prevention 26, 170–176. doi:10.1136/injuryprev-2019-043426.

Stavrinos, D., Byington, K.W., Schwebel, D.C., 2011. Distracted walking: Cell phones increase injury risk for college pedestrians. Journal of Safety Research 42, 101–107. doi:10.1016/j.jsr.2011.01.004.

Sun, H., Hua, C., Luo, Y., 2004. A multi-stage classifier based algorithm of pedestrian detection in night with a near infrared camera in a moving car. Proceedings - Third International Conference on Image and Graphics , 120–123doi:10.1109/icig.2004.11.

Supernak, J., Verma, V., Supernak, I., 2013. Pedestrian Countdown Signals: What Impact on Safe Crossing? Open Journal of Civil Engineering 03, 39–45. doi:10.4236/ojce.2013.33b007.

Taherdoost, H., 2017. Determining sample size; How to calculate survey sample size. International Journal of Economics and Management Systems 2, 237–239.

Theil, H., 1958. Economic Forecasts and Policy. Amsterdam: North Holland.

Theil, H., 1966. Applied Economic Forecasting. Chicago: Rand McNally.

Thompson, S., Monsere, C., Figliozzi, M., Koonce, P., Obery, G., 2013. Bicycle-specific traffic signals. Transportation Research Record , 1–9doi:10.3141/2387-01.

Tiwari, G., Bangdiwala, S., Saraswat, A., Gaurav, S., 2007. Survival analysis: Pedestrian risk exposure at signalized intersections. Transportation Research Part F: Traffic Psychology and Behaviour 10, 77–89. doi:10.1016/j.trf.2006.06.002.

Tom, A., Granié, M.A., 2011. Gender differences in pedestrian rule compliance and visual search at signalized and unsignalized crossroads. Accident Analysis and Prevention 43, 1794–1801. doi:10.1016/j.aap.2011.04.012.

Tyler T.R., 1990. Why People Obey the Law. New Haven: Yale University Press .

Van Houten, R., Ellis, R., Kim, J.L., 2007. Effects of Various Minimum Green Times on Percentage of Pedestrians Waiting for Midblock "Walk" Signal. Transportation Research Record: Journal of the Transportation Research Board 2002, 78–83. doi:10.3141/2002-10.

Vasudevan, V., Mehta, M., Dutta, B., 2020. Pedestrian temporal gap acceptance behavior at unsignalized intersections in Kanpur, India. Transportation Research Part F: Traffic Psychology and Behaviour 74, 95–103. doi:10.1016/j.trf.2020.08.010.

Violano, P., Roney, L., Bechtel, K., 2015. The incidence of pedestrian distraction at urban intersections after implementation of a Streets Smarts campaign. Injury Epidemiology 2, 18. doi:10.1186/s40621-015-0050-7.

Vujanić, M., Pešić, D., Antić, B., Smailović, E., 2014. Pedestrian Risk At The Signalized Pedestrian Crossing Equipped With Countdown Display. International

Journal For Traffic And Transport Engineering 4, 52–61. doi:10.7708/ijtte.2014.4(1).04.

Walker, E.J., Lanthier, S.N., Risko, E.F., Kingstone, A., 2012. The effects of personal music devices on pedestrian behaviour. Safety Science 50, 123–128. doi:10.1016/j.ssci.2011.07.011.

Wang, W., Guo, H., Gao, Z., Bubb, H., 2011. Individual differences of pedestrian behaviour in midblock crosswalk and intersection. International Journal of Crashworthiness 16, 1–9. doi:10.1080/13588265.2010.491715.

Wanjing, M., Liao, D., Bai, Y., 2015. Empirical analysis of countdown signals on pedestrian behaviour. Transport 168, 15–22. URL: http://dx.doi.org/10.1680/tran.11.00078.

Washington, S., Karlaftis, M., Mannering., F., 2011. Statistical and Econometric Methods for Transportation Data Analysis. 2nd ed., CRC Press, Boca Raton, Fla.

Wells, H.L., McClure, L.A., Porter, B.E., Schwebel, D.C., 2017. Distracted Pedestrian Behavior on two Urban College Campuses. Journal of Community Health 43, 96–102. doi:10.1007/s10900-017-0392-x.

World Bank Group, 2020. Labor Force Participation Rate, Female (% of Female Population Ages 15+) (Modeled ILO Estimate) – India. URL: https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?locations=IN.

Xu, Y., Li, Y., Zhang, F., 2013. Pedestrians' intention to jaywalk: Automatic or planned? A study based on a dual-process model in China. Accident Analysis and Prevention 50, 811–819. doi:10.1016/j.aap.2012.07.007.

Yagil, D., 2000. Beliefs, motives and situational factors related to pedestrians' self-reported behavior at signal-controlled crossings. Transportation Research Part F: Traffic Psychology and Behaviour 3, 1–13. doi:10.1016/S1369-8478(00)00004-8.

Yang, C.Y., Najm, W.G., 2007. Examining driver behavior using data gathered from red light photo enforcement cameras. Journal of Safety Research 38, 311–321. doi:10.1016/j.jsr.2007.01.008.

Yang, X., Abdel-Aty, M., Huan, M., Peng, Y., Gao, Z., 2015. An accelerated failure time model for investigating pedestrian crossing behavior and waiting times at signalized intersections. Accident Analysis and Prevention 82, 154–162. doi:10.1016/j.aap.2015.04.036.

Yang, Y., Sun, J., 2013. Study on pedestrian red-time crossing behavior. Transportation Research Record , 117–124doi:10.3141/2393-13.

Yannis, G., Papadimitriou, E., Theofilatos, A., 2013. Pedestrian gap acceptance for mid-block street crossing. Transportation Planning and Technology 36, 450–462. doi:10.1080/03081060.2013.818274.

Zegeer, C.V., Bushell, M., 2012. Pedestrian crash trends and potential countermeasures from around the world. Accident Analysis and Prevention 44, 3–11. doi:10.1016/j.aap.2010.12.007.

Zhang, L., Zhang, C., Shang, L., 2016. Sensation-seeking and domain-specific risk-taking behavior among adolescents: Risk perceptions and expected benefits as mediators. Personality and Individual Differences 101, 299–305. doi:10.1016/j.paid.2016.06.002.

Zhang H. and Zhang C. and Chen F. and We IY., 2017. The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections, in: 4th International Conference on Transportation Information and Safety (ICTIS), pp. 197–205.

Zhao, J., Malenje, J.O., Tang, Y., Han, Y., 2019. Gap acceptance probability model for pedestrians at unsignalized mid-block crosswalks based on logistic regression. Accident Analysis and Prevention 129, 76–83. doi:10.1016/j.aap.2019.05.012.

Zhou, R., Horrey, W.J., Yu, R., 2009. The effect of conformity tendency on pedestrians' road-crossing intentions in China: An application of the theory of planned behavior. Accident Analysis and Prevention 41, 491–497. doi:10.1016/j.aap.2009.01.007.

Zhou, Z., Liu, S., Xu, W., Pu, Z., 2019. Impacts of mobile phone distractions on pedestrian crossing behavior at signalized intersections: An observational study in China. Advances in Mechanical Engineering 11, 1–8. doi:10.1177/1687814019841838.

Zhu, H., Almukdad, A., Iryo-Asano, M., Alhajyaseen, W.K., Nakamura, H., Zhang, X., 2021. A novel agent-based framework for evaluating pedestrian safety at unsignalized mid-block crosswalks. Accident Analysis & Prevention 159, 106288. doi:10.1016/j.aap.2021.106288.

Zhuang, X., Wu, C., 2011. Pedestrians' crossing behaviors and safety at unmarked roadway in China. Accident Analysis & Prevention 43, 1927–1936. doi:10.1016/j.aap.2011.05.005.

Zhuang, X., Wu, C., 2012. The safety margin and perceived safety of pedestrians at unmarked roadway. Transportation Research Part F: Traffic Psychology and

Behaviour 15, 119–131. URL: https://linkinghub.elsevier.com/retrieve/pii/S1369847811001045, doi:10.1016/j.trf.2011.11.005.

List of Publications

Journals

1. **Raoniar, R.**, Singh, S., Pathak, A. and Maurya, A. K., 2022. Evaluation of Pedestrian-Vehicle Interaction at Signalised Intersection Crosswalks: A Random Intercept Regression Approach. [*Under Preparation*].

2. **Raoniar, R.** and Maurya, A. K., 2022. Digital and Social Distractions' Impact on Pedestrian Road Crossing Behaviour at Signalized Intersection Crosswalks, [*Under Second Review*].

3. **Raoniar, R.**, Maqbool, S., Pathak, A., Chugh, M. and Maurya, A. K., 2022. Hazard-Based Duration Approach for Understanding Pedestrian Crossing Risk Exposure at Signalised Intersection Crosswalks - A Case Study of Kolkata, India, Transportation Research Part-F, vol. 85, pp. 47-68, doi: https://doi.org/10.1016/j.trf.2021.12.015.

4. **Raoniar, R.** and Maurya, A. K., 2022. Pedestrian Red-Light Violation at Signalized Intersection Crosswalks: Influence of Social and Non-Social Factors, Safety Science, vol. 147, doi: https://doi.org/10.1016/j.ssci.2021.105583

5. **Raoniar, R.** and Maurya, A. K., 2021. Pedestrian Crossing Behaviour at Signalized Intersection Crosswalks: An Observational Study of Factors Influencing Pedestrian Walking Speed, Safety Margin, and Violation, Journal of the Eastern Asia Society for Transportation Studies, vol. 13, doi: https://doi.org/10.11175/easts.14.1456.

6. Das, S., Boruah, A., Banerjee, A., **Raoniar, R.**, Nama, S. and Maurya, A.K. 2021. Impact of COVID-19: A radical modal shift from public to private transport mode, Transport Policy, vol. 109, doi: https://doi.org/10.1016/j.tranpol.2021.05.005.

7. Banerjee, A., **Raoniar, R.** and Maurya, A.K., 2020. Pedestrian overpass utilization modelling based on mobility friction, safety and security, and connectivity using machine learning techniques. Soft Computing, vol. 24, 17467–17493, doi: https://doi.org/10.1007/s00500-020-05277-w

8. **Raoniar R**, Das T, Banerjee A and Maurya, A.K., 2019. The Parents' Role in School Mode Choice for their Children: A Case Study in Guwahati, Journal of the Eastern Asia Society for Transportation Studies, vol. 13, 775-94, doi: https://doi.org/10.11175/easts.13.775

Presentations and Proceedings in International/National Conferences

1. Banerjee, A., **Raoniar, R.** and Maurya, A. K., 2021. Study of Factors Impacting Safety-Security and Mobility Friction on the Choice of Pedestrians in Using Skywalk Facilities through Soft Computing Approaches, Journal of the Eastern Asia Society for Transportation Studies, vol. 13, http://www.easts.info/on-line/proceedings/vol.13/head.htm.

2. **Raoniar, R.** and Maurya, A.K., 2021. Hazard-Based Duration Approach: Pedestrian Crossing Risk Exposure at Signalised Intersections in Kolkata, India, 6th CTRG (ID-24), India.

3. **Raoniar, R.** and Maurya, A.K., 2021. Impact of Different Distractions on Pedestrian Road Crossing Behaviour at Signalized Intersection Crosswalks, [ID: TRBAM-21-01200 (Poster presentation)], 100th Annual Meeting of Transportation Research Board, Washington, DC, United States.
   https://annualmeeting.mytrb.org/OnlineProgram/Details/15701

4. **Raoniar, R.**, Maqbool, S. and Maurya, A.K., 2020. Factors Influencing Pedestrian Walking Speed and Safety Margin at Signalized Intersection Crosswalks, 13th TPMDC, Paper ID 86, IIT Bombay.
   https://www.civil.iitb.ac.in/~tse/tpmdc_web/docs/TPMDC_2020_Session_Details.pdf

5. Banerjee, A., **Raoniar, R.** and Maurya, A.K., 2020. Identification of factors influencing the pedestrian walking speed over elevated facilities using Deep Neural Network, 13th TPMDC, Paper ID 136, IIT Bombay.
   https://www.civil.iitb.ac.in/~tse/tpmdc_web/docs/TPMDC_2020_Session_Details.pdf

6. **Raoniar, R.** and Maurya, A.K., 2020. Predicting Pedestrian's Red Light Violation (RLV) Behavior at Signalized Intersection Crosswalks using Machine Learning Techniques, ASCEic (ID: AIC2020-10-932).