RECENT TREND IN CIVIL ENGINEERING AND BUILT ENVIRONMENT
e-ISSN: 2773-5303

RTCEBE

Vol. 6 No. 1 (2025) 256-268
https://publisher.uthm.edu.my/periodicals/index.php/rtcebe

Influence of Traffic Conditions on Pedestrian Crossing Decision at Urban Street Crosswalk

Fakhrul Amer Fahami[1], Basil David Daniel [2]

[1,2] *Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia (UTHM), 86400 Parit Raja, Batu Pahat, Johor, MALAYSIA*

*Corresponding Author: basil@uthm.edu.my
DOI: https://doi.org/10.30880/rtcebe.2025.06.01.027

Article Info

Received: 08 January 2024
Accepted: 15 April 2024
Available online: 6 May 2025

Keywords

Pedestrian Crossing Decision, Raised crosswalks, zebra crossings, Pedestrian Safety, Pedestrian behaviour.

Abstract

This study investigates pedestrian decision-making at urban road-crossing using a mixed-methods approach, incorporating surveys, on-site observations, Exploratory Factor Analysis (EFA), and Binary Logistic Regression. Key findings reveal physical and psychosocial factors, such as the type of vehicle, presence of pedestrians, and traffic volume, significantly influence pedestrian choices. EFA identifies two major elements impacting decisions, while Binary Logistic Regression verifies their predictive potential. High traffic volume deters crossings, suggesting a need for congestion reduction, while positive coefficients for vehicle speeding, yielding behavior, and raised crosswalks encourage crossings. The study recommends evidence-based interventions, including traffic calming measures and awareness campaigns, to enhance pedestrian safety and traffic flow.

1. Introduction

Pedestrian safety in urban areas, especially at street crosswalks, is a crucial concern, and this research investigates the intricate relationship between pedestrian decision-making and traffic conditions (Zhang et al., 2018). The findings underscore the significant impact of various factors on pedestrians' choices, shedding light on the complexities of navigating urban roadways. The context emphasizes the global challenge of pedestrian accidents, with a specific focus on the escalating road safety issues in Malaysia, where pedestrian-related fatalities have been on the rise (Nasrudin et al., 2020). The observed trends reveal a pressing need for in-depth exploration and understanding of how traffic conditions influence pedestrians' decisions.

The identified gap in existing research highlights the limited comprehension of the specific factors shaping pedestrian choices in urban environments. This research aims to bridge this gap by providing a comprehensive analysis of the intricate dynamics between pedestrians and traffic. The objectives of the study are twofold: firstly, to identify and comprehend the traffic-related factors significantly impacting pedestrian crossing decisions, and secondly, to develop a pedestrian crossing decision model based on these traffic conditions.

The scope of this study focuses on the urban environment of Batu Pahat, utilizing theoretical frameworks and advanced statistical analysis to investigate pedestrian decision-making at selected crosswalks (Amini et al., 2019). The research integrates a mixed-methods approach, combining survey responses, on-site video observations, and sophisticated statistical tools such as Exploratory Factor Analysis (EFA) and Binary Logistic Regression. These methodologies aim to provide a nuanced understanding of the multifaceted factors influencing pedestrian choices.

The significance of this research is rooted in its potential to contribute evidence-based interventions and policies that enhance pedestrian safety in urban settings. The findings presented in subsequent chapters will offer insights into how specific traffic-related factors, such as traffic volume, vehicle speeds, and the presence of other pedestrians or vehicles, impact pedestrian decision-making. Ultimately, this research endeavours to inform the development of practical solutions, contributing to safer and more pedestrian-friendly urban environments.

2. Common Factors That Govern Pedestrian Crossing Decision

2.1 Pedestrian Gap Acceptance

From a pedestrian perspective, the time or distance gap with the closest oncoming vehicle must be long enough for safely crossing the road (Velasco et al., 2017). The required time for pedestrians to cross, known as the "critical safe gap," is computed based on the crossing length and pedestrian crossing speed. Pedestrians need larger gaps for safe crossing maneuvers, which may result in longer waiting time and increased risk-taking behavior to accept shorter gaps. The likelihood of pedestrians accepting shorter gaps also declines when they are in a group compared to individual pedestrians (Amini et al., 2019). At unsignalized crossings, pedestrians seek rolling gaps in high traffic volume, meaning that they predict where a gap will be available in the next lane and do not wait for all lanes to be clear (Camara et al., 2018). The study of gap acceptance behavior at urban crosswalks is crucial for understanding pedestrian behavior and improving traffic management (Masoud et al., 2016). Sun et al. (2003) used the probabilistic model and the binary logistic regression model, respectively, to describe pedestrian gap acceptance behaviors and driver yielding behaviors at mid-block locations.

2.2 Vehicle Yielding Behavior

Driver behavior is influenced by personal characteristics, vehicle type, and the traffic ecosystem (Goddard et al., 2015). The interaction between vehicles and pedestrians is characterized by subtle non-verbal communication of intent. Both parties need to accurately perceive each other's actions and intentions, relying on cues like hand signals, flashing lights, or eye contact (Obeid et al., 2017; Sucha et al., 2017). Higher vehicle speeds and platooning reduce the likelihood of drivers yielding to pedestrians (Bertulis & Dulaski, 2014; Wang et al., 2016). Studies have shown that bold, assertive, and aggressive crossing behavior can prompt drivers to yield (Tasic et al., 2017).

2.3 Pedestrian Group Size

The size of pedestrian platoons significantly influences both pedestrian crossing behavior and driver yielding, with the likelihood of yielding doubling when the number of pedestrians waiting at crosswalks increases from three to six persons, as larger groups capture more attention from drivers (Sucha et al., 2017). It has also been found that the possibility of a group of four pedestrians continuing to traverse the crosswalk is 70% higher than an individual pedestrian. This also enhances the safety of pedestrians as pedestrians in groups are more detectable (Amini et al., 2019). Due to the frequency of crossing, smaller pedestrian platoons cause more traffic interruptions and higher cumulative delay, and consequently driver inconvenience compared to bigger platoons that cross at once (Malenje et al., 2018).

2.4 Size of Approaching Vehicle

The size of vehicles can influence the behavior of both pedestrians and drivers and is an important factor for modelling road users' behaviors (Sun et al., 2015). when the oncoming vehicle is a lorry or bus instead of a passenger car. The number of attempts by pedestrians to cross also declines if the oncoming vehicle is a heavy vehicle, such as a bus or a coach (Hamed, 2001). The probability of drivers reacting is also 30% lower among heavy vehicle drivers than car drivers (Amini et al., 2019).

2.5 Traffic Volume

Traffic density is a factor that pedestrians consider when making the crossing decision; high traffic density is considered a risky situation (Papadimitriou et al., 2016). Pedestrians also make more attempts to cross the roadway in heavy traffic volume and may be more likely to accept shorter gaps from individual vehicles than platoons. On the other hand, the driver's tendency to give priority to pedestrians declines in high traffic volume. Moreover, the possibility of the lead driver reacting decreases with the increasing number of approaching vehicles. In a group of five vehicles, the probability of reacting can decrease by almost 40% more than with only one vehicle (Amini et al., 2019).

Recent Trend in Civil Engineering and Built Environment Vol. 6 No. 1 (2025) p. 256-268

3. Methodology

This section outlines the research methodology, incorporating both survey and on-site observation methods. Batu Pahat is the chosen location for its relevance to urban pedestrian behavior. The survey method is integral, involving literature review, questionnaire design, and rigorous analysis, while on-site observation utilizes video recording for real-time data collection.

3.1 Survey

3.1.1 Questionnaire Design

A structured questionnaire, widely acknowledged as a valuable tool for research data collection (Bird, 2009), was designed to gather information on pedestrians' crossing decisions and perceptions of identified factors. This questionnaire comprised two main sections: the demographic section, obtaining basic participant information such as age and gender to establish socio-demographic characteristics, and the section focusing on factors influencing pedestrian crossing decisions. In this section, participants rated the influence of eight identified factors, including type of vehicle, pedestrian density, traffic volume, gap acceptance, presence of zebra crossing (raised or standard), driver speeding, driver yielding, and crossing distance, using a Likert scale ranging from strongly disagree to strongly agree, capturing both quantitative and qualitative insights into pedestrians' behavior, attitudes, and perceptions.

3.1.2 Reliability Test Using Cronbach's Alpha Coefficient

Before the questionnaire is distributed randomly to the respondents, the validity, reliability, and acceptability of the question in the questionnaire should be measured and tested (Williams, 2003). Cronbach's Alpha coefficients are used to determine the internal consistency of the questions. The alpha coefficients can be a value from 0-1. The alpha coefficients that exceed 0.7 are considered acceptable based on Table 1 (Gliem & Gliem, 2003).

Table 1 Cronbach's Alpha Coefficients (Gliem & Gliem, 2003)

Cronbach's Alpha	Internal Consistency
$\alpha \geq 0.9$	Excellent
$0.8 \leq \alpha \leq 0.9$	Good
$0.7 \leq \alpha \leq 0.8$	Acceptable
$0.6 \leq \alpha \leq 0.7$	Questionable
$0.5 \leq \alpha \leq 0.6$	Poor
$\alpha < 0.5$	Unacceptable

Table 2 Reliability Statistics of Pilot Study

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.808	.803	10

The Alpha Cronbach coefficient for this pilot research is 0.808, which is deemed good internal consistency according to Table 2. This questionnaire's items are all appropriate for this investigation.

3.2 On-Site Observation

3.2.1 Video Recording

The video recordings aimed to capture real-life instances of pedestrians' crossing decisions and the corresponding traffic conditions. This visual data provided rich qualitative information, allowing for a detailed analysis of pedestrians' behaviors, interactions with traffic, and responses to various traffic conditions.

3.3 Data Analysis

3.3.1 Exploratory Factor Analysis

Exploratory factor analysis (EFA) was conducted to explore the underlying factor structure of the identified factors influencing pedestrian crossing decisions. EFA is a statistical technique that allows for the identification of latent constructs or underlying dimensions in a set of observed variables. In this case, EFA helped to determine the interrelationships among the eight factors and identify the key dimensions that explain pedestrians' crossing decisions. Factor loadings, eigenvalues, and the percentage of variance explained were examined to determine the strength and significance of each factor. Factor loadings represent the correlation between the observed variables and the latent factors. Higher factor loadings indicate a stronger association between a specific factor and pedestrians' crossing decisions.

3.3.2 Logistic Binary Regression

To examine the relationship between the identified influential factors and pedestrians crossing decisions, binary logistic regression analysis was performed using statistical software, such as SPSS. Binary logistic regression is a statistical technique that allows for the analysis of the relationship between a binary dependent variable (pedestrians' crossing decisions) and one or more independent variables (the identified influential factors). It provides insights into the odds ratios, statistical significance, and the direction of the relationships between the factors and pedestrian crossing behavior. The dependent variable for the binary logistic regression analysis was the pedestrians' crossing decision, coded as 0 for crossing and 1 for not crossing. The independent variables were the influential factors identified through the survey phase. Statistical significance was assessed using p-values, which indicated whether the relationship between each factor and the pedestrian crossing decision was statistically significant. A significance level of $\alpha = 0.05$ was used to determine the presence of statistically significant associations. The regression equation takes following form:

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \ldots + \beta nXn \tag{1}$$

Where Y is a dependent variable in the study. Each X1, X2, ... Xn is a particular independent variable and the β0, β1, β2, ... βn, are coefficients of regression of independent variables in the study.

4. Result And Analysis

4.1 Respondents' Acceptance of Physical and Psychosocial Factors

The mean scores, clearly depicted in table 3, provide numerical insights into the perceived significance of each factor. Notably, the mean values across all factors are strikingly high, with Type of Vehicle at 4.29, Presence of Pedestrians at 4.00, Traffic Volume at 3.99, Acceptance Gap at 4.09, Presence and Design of Crosswalk at 4.18, Driver Behaviour - Speed at 4.15, Driver Behaviour - Yielding at 4.14, Number of Road Lanes (Crossing Distance) at 4.12, and Presence and Design of Crosswalk at 4.11. These high mean values underscore the considerable importance attributed by respondents to each of these factors. Additionally, the generally low standard deviations, ranging from 0.631 to 0.723, suggest a high degree of consensus among participants, emphasizing the consistent influence of these factors in shaping pedestrians' decisions to cross a road.

Table 3 Statistic Mean and Standard Deviation of Physical and Psychosocial Factors

Factors	*Type of Vehicle*	*Presence of Pedestrians*	*Traffic Volume*	*Acceptance Gap*	*Presence and Design of Crosswalk*	*Driver Behaviour - Speed*	*Driver Behaviour - Yielding*	*Number of Road Lanes (Crossing Distance)*	*Presence and Design of Crosswalk*
Mean	4.29	4.00	3.99	4.09	4.18	4.15	4.14	4.12	4.11
Std. Deviation	.647	.665	.723	.707	.645	.655	.637	.631	.660

4.2 Video Recording Analysis

The collected data in Table 4 was methodically categorized into the subsequent variable designations, encompassing key aspects of pedestrian action.

Table 4 Video Recording Analysis

Variable Name	Video Recording Observation	Analysis of Recording
Type of Vehicle	It was observed that pedestrians often hesitated when crossing in the proximity of larger vehicles, such as trucks.	This shows that vehicle type may affect their decision-making, might be due to perceived difficulties manoeuvring larger or slower cars.
Presence of Larger Number of Pedestrians	In crowded scenarios captured in the video recordings, pedestrians tended to display increased confidence when crossing in groups.	The presence of a larger number of pedestrians appeared to positively influence their decision to cross.
Traffic Volume	Video analysis revealed that pedestrians exhibited a trend of increased caution and longer waiting times in areas with higher traffic volumes	The number of vehicles on the road appeared to influence their decision-making, with a notable impact on crossing patterns.
Gap Acceptance between Vehicles and Crossing Area	Pedestrians consistently demonstrated a tendency to wait for a safe distance between vehicles before initiating a road crossing.	This behaviour suggests a conscious consideration of the safe distance factor in their decision-making process.
Driver Behaviour (speeding)	Pedestrians exhibited heightened caution and increased wait times in the presence of fast-moving vehicles.	The video recordings suggest that driver behaviour, particularly high speeds, influences pedestrians to adopt more conservative crossing behaviours.
Driver Behaviour (yielding)	Pedestrians displayed a clear trend of increased confidence and swifter crossings when drivers yielded the right of way.	The observed instances emphasize the positive impact of driver behaviour on pedestrians' decision-making processes.

Number of Road Lanes (Crossing Distance)	In areas with multiple road lanes, pedestrians consistently exhibited cautious decision, carefully assessing, and strategizing their crossing.	The number of road lanes emerged as a significant factor influencing pedestrians' decision-making, particularly in complex traffic configurations.
Presence of Raised Zebra Crossing	Instances in the video footage depicted a notable increase in pedestrians utilizing zebra crossings.	The presence of such infrastructure proved to positively affect their decision to cross at specified places, highlighting the importance of well-designed crossings.
Presence of Zebra Crossing	Video analysis demonstrated a clear inclination among pedestrians to utilize standard zebra crossings.	The marked crosswalks influenced their decision to cross at designated locations, emphasizing the importance of recognizable pedestrian infrastructure in shaping crossing decision.

4.3 Result Analysis for Exploratory Factor Analysis (EFA)

4.3.1 Correlation Matrix[a]

Table 4 reveals interesting patterns among the survey questions from factor 1 to factor 9 (QS1-QS9), providing insights into potential latent constructs influencing pedestrian decision-making. Notably, there is a moderate positive correlation of 0.481 between Factor 1 and Factor 8, indicating consistent responses to questions related to factors like the presence and design of crosswalks and driver behavior yielding. This suggests that pedestrians who consider one of these factors might also consider the other.

Additionally, the positive correlation between Factor 4 (Acceptance Gap) and Factor 8 (Number of Road Lanes) with coefficients of 0.507 and 0.452, respectively, highlights the interrelatedness of factors related to traffic conditions. Pedestrians who prioritize factors such as the acceptance gap between vehicles and crossing areas might also be influenced by the number of road lanes when deciding to cross. The generally positive correlations across various pairs of questions suggest that pedestrians may weigh multiple factors simultaneously when making road-crossing decisions.

Table 4 Correlation Matrix[a] for EFA Analysis

Correlation Matrix[a]

		QS1	QS2	QS3	QS4	QS5	QS6	QS7	QS8	QS9
Correlation	QS1	1.000	.210	.264	.407	.299	.312	.263	.432	.270
	QS2	.210	1.000	.335	.417	.293	.323	.296	.419	.481
	QS3	.264	.335	1.000	.434	.348	.332	.232	.400	.213
	QS4	.407	.417	.434	1.000	.386	.373	.396	.507	.347
	QS5	.299	.293	.348	.386	1.000	.413	.292	.368	.299
	QS6	.312	.323	.332	.373	.413	1.000	.501	.479	.358
	QS7	.263	.296	.232	.396	.292	.501	1.000	.469	.286
	QS8	.432	.419	.400	.507	.368	.479	.469	1.000	.452
	QS9	.270	.481	.213	.347	.299	.358	.286	.452	1.000

a. Determinant = .076

Recent Trend in Civil Engineering and Built Environment Vol. 6 No. 1 (2025) p. 256-268

4.3.2 KMO and Bartlett's Test

From the Table 5, the results from the Kaiser-Meyer-Olkin (KMO) measure (0.874) and Bartlett's Test of Sphericity (Approx. Chi-Square = 501.853, df = 36, Sig. < 0.001) collectively indicate that the dataset is highly suitable for exploratory factor analysis (EFA). The elevated KMO measure suggests excellent sampling adequacy, affirming that the variables share significant common variance. Concurrently, the significant outcome of Bartlett's Test demonstrates non-random correlations between variables, providing compelling evidence for the dataset's appropriateness for factor analysis. These findings collectively support the robustness of the dataset in uncovering latent factors influencing pedestrian decision-making regarding road crossings in urban areas.

Table 5 KMO and Bartlett's Test for EFA Analysis

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.874
Bartlett's Test of Sphericity	Approx. Chi-Square	501.853
	df	36
	Sig.	.000

4.3.3 Total Variance Explained

The discussion on Eigenvalues and variance distribution in the Exploratory Factor Analysis (EFA) indicates that the first three components (Component 1, Component 2, and Component 3) are particularly significant in capturing the underlying structure of the dataset. The Eigenvalue analysis in table 6 reveals that the first component dominates, representing 43.339% of the total variance, and subsequent components contribute progressively less. The cumulative variance reaches 62.834% by the third component, highlighting a balanced trade-off between model complexity and explanatory power. This trend is consistent in both the Extraction Sums of Squared Loadings and Rotation Sums of Squared Loadings tables, where the first three components consistently capture a substantial portion of the variance. The diminishing returns in variance contribution with additional components underscore the trade-off between extracting more factors and maintaining simplicity in the model. Therefore, it can be inferred that, based on these analyses, components 1, 2, and 3 (factor 1,2, and 3) are deemed more significant in explaining the patterns and relationships within the dataset during the EFA.

Table 6 Table Total Variance Explained

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.900	43.339	43.339	3.900	43.339	43.339	1.031	11.456	11.456
2	.899	9.994	53.333	.899	9.994	53.333	1.026	11.401	22.857
3	.855	9.501	62.834	.855	9.501	62.834	1.025	11.393	34.250
4	.773	8.593	71.427	.773	8.593	71.427	1.018	11.315	45.565
5	.691	7.682	79.109	.691	7.682	79.109	1.017	11.296	56.860
6	.553	6.141	85.250	.553	6.141	85.250	1.010	11.226	68.087
7	.476	5.291	90.541	.476	5.291	90.541	.985	10.945	79.032
8	.438	4.864	95.406	.438	4.864	95.406	.968	10.756	89.788
9	.413	4.594	100.000	.413	4.594	100.000	.919	10.212	100.000

Extraction Method: Principal Component Analysis.

4.3.4 Scree Plot

Figure 1 displays the Scree Plot Graph, a crucial tool in exploratory factor analysis (EFA). The Scree Plot offers valuable insights into determining the optimal number of components for our dataset. The eigenvalues show a sharp decline after the second component, creating a distinct elbow in the plot. This suggests that the first two components represent a substantial proportion of the total variance, while subsequent components contribute less significantly. In this context, a two-factor solution seems appropriate, aligning with the point where the eigenvalues plateau. This choice ensures a concise model that captures the essential patterns in the data without unnecessary complexity. In this context, the scree plot strongly advocates for a two-factor solution, aligning with the notion that the dataset's underlying structure is best characterized by these two key factors. This information guides our decision on the number of factors to retain, facilitating a straightforward and understandable model.

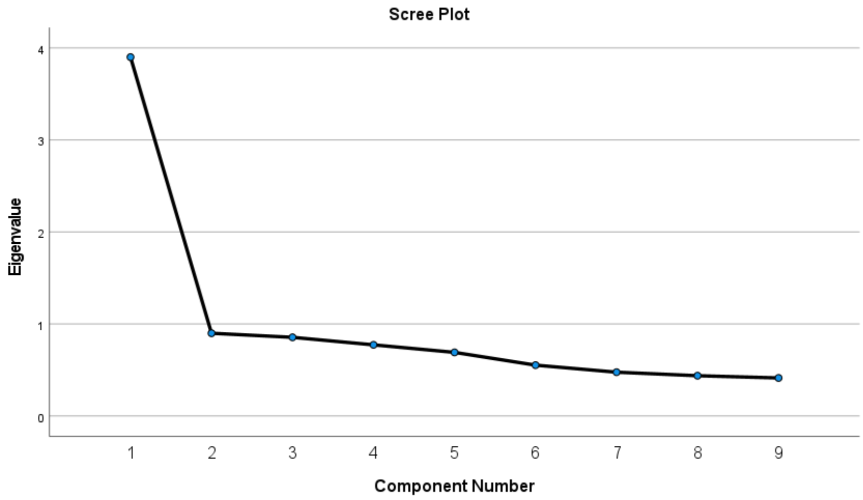

Fig 1 Scree Plot Graph

4.3.5 Rotated Component Matrix[a]

Table 7 shows the rotated component matrixa resulting from the Principal Component Analysis with varimax rotation and Kaiser normalization reveals distinct factor loadings for each variable across the identified components. Notably, variable QS1 (Type of Vehicle) exhibits a high loading of 0.952 on Component 1, along with substantial contributions from QS4 (Gap Acceptance) is 0.177 and QS8 (Crossing Distance) is 0.193 to the same component. In Component 2, QS2 (Pedestrian Density) stands out with a loading of 0.928, accompanied by notable loadings from QS6 (Vehicle Speed) is 0.226 and QS8 (Crossing Distance) is 0.202. Variable QS3 (Traffic Volume) plays a central role in Component 3, displaying a high loading of 0.945, while Component 4 is characterized by a dominant loading from QS4 (Gap Acceptance) which is 0.894. Additionally, QS5 (Presence of Raised Zebra Crossing) exhibits a substantial loading of 0.941 in Component 5, while QS7 (Driver Yielding) takes a prominent position in Component 7 with a loading of 0.932. Component 6 is influenced by high loadings from QS2 (Pedestrian Density) is 0.858 and QS7 (Driver Yielding) is 0.906. Lastly, Component 9 is marked by a noteworthy loading of 0.933 from QS9 (Presence of Zebra Crossing).

Table 7 Table of Rotated Component Matrix[a]

Rotated Component Matrix[a]

	Component								
	1	2	3	4	5	6	7	8	9
QS1	.952	.086	.092	.093	.110	.055	.102	.149	.148
QS2	.058	.102	.133	.215	.101	.928	.101	.150	.133
QS3	.094	.066	.945	.055	.137	.127	.113	.162	.130
QS4	.177	.157	.189	.122	.150	.167	.116	.894	.173

QS5	.114	.098	.139	.106	.941	.098	.156	.131	.106
QS6	.116	.226	.127	.135	.172	.108	.906	.112	.160
QS7	.090	.932	.067	.094	.100	.101	.208	.140	.163
QS8	.193	.202	.165	.192	.131	.161	.182	.189	.864
QS9	.099	.094	.056	.933	.107	.212	.125	.109	.155

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 6 iterations.

4.4 Result Analysis for Binary Logistic Regression

4.4.1 Dependent Variable

Table 8 shows the dependent variable was encoded to facilitate the modelling process. The original values of the dependent variable were "YA" (Cross) and "TIDAK" (Not Cross), which were recoded into internal values for computational purposes. The internal coding assigned the value of 0 to "YA" and 1 to "TIDAK." This coding scheme is essential for interpreting the logistic regression results. In the context of this study, the internal value of 0 indicates the absence of the event or the positive outcome (crossing the road), while the internal value of 1 signifies the presence of the event or the negative outcome (not crossing the road). Throughout the subsequent analysis, references to the dependent variable will be made using these internal values, providing a standardized approach to understanding the logistic regression model outputs.

Table 8 Table of Dependent Variable

Dependent Variable Encoding

Original Value	Internal Value
YA	0
TIDAK	1

4.4.2 Omnibus Tests of Model Coefficients

The Omnibus Tests of Model Coefficients, as indicated in table 9, play a crucial role in assessing the overall significance of our logistic regression model. The Chi-square statistic, with a value of 20.119 and 9 degrees of freedom for each step (Step, Block, and Model), reveals that there is a significant association between the predictor variables and the binary outcome variable. The associated p-value of .017 indicates that the observed results are statistically significant at a conventional significance level of 0.05. This suggests that the model or its components collectively contribute significantly to explaining the variance in the dependent variable. As such, we have evidence to support the overall effectiveness of our logistic regression model in predicting the outcome variable.

Table 9 Table of Omnibus Tests of Model Coefficients

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	20.119	9	.017
	Block	20.119	9	.017
	Model	20.119	9	.017

4.4.3 Cox & Snell R Square and Nagelkerke R Square

The percentage of the total variation of the dependent variable that is explained by the model is measured by the Cox & Snell R Square and the Nagelkerke R Square. In Table 10 the shows, shows the explained variation in the dependent variable by the model ranges from 9.6% to 40.5% and correctly classified 98% of cases. The -2 Log Likelihood value of 33.778 indicates the overall fit of the model to the data, with lower values suggesting a better fit. The Cox & Snell R Square, standing at 0.096, represents the proportion of variability in the dependent variable explained by the model. Additionally, the Nagelkerke R Square, with a value of 0.405, is an adjusted measure of the model's explanatory power. These values collectively suggest that the logistic regression model has a moderate level of explanatory ability, capturing a substantial portion of the variability in the dependent variable. It is crucial to interpret these metrics in conjunction with other diagnostic measures to comprehensively assess the model's performance and predictive capabilities. The termination of estimation at iteration 8, due to minimal changes in parameter estimates, indicates stability in the model and supports the reliability of the reported statistics.

Table 10 Table of Model Summary

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	33.778[a]	.096	.405

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

4.4.4 Hosmer and Lemeshow Test

The Hosmer and Lemeshow test are statistical tests for goodness of fit for the binary logistic regression model. It indicates a good fit if the significance value is greater than 0.05 (P>0.05). The Hosmer and Lemeshow Test results in Table 11 reveal a Chi-square statistic of 32.418 with 8 degrees of freedom, yielding a highly significant p-value of less than 0.001. This test assesses the goodness of fit for the binary logistic regression model, specifically examining the agreement between observed and expected values across different levels of the dependent variable. The obtained p-value suggests that there is a significant difference between the observed and expected frequencies, indicating a lack of fit for the model. Further scrutiny and potential model adjustments may be necessary to improve the overall fit and predictive accuracy. The small p-value indicates that the model does not fit the observed data well.

Table 11 Table of Hosmer and Lemeshow Test

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	32.418	8	.000

4.4.5 Binary Logistic Regression Model Estimation

Table 12 shows the logistic regression analysis revealing varying degrees of significance among the predictor variables. The logistic regression model provides insights into the factors influencing the likelihood of crossing decision among pedestrians. Examining the coefficients and odds ratios from the provided table, Q1 (Type of Vehicle) has a coefficient of -0.188, indicating a lower likelihood for the specified vehicle type compared to the reference category. Q3 (Traffic Volume) has a notable negative coefficient of -1.902, suggesting a significantly reduced likelihood of higher traffic volume. Conversely, Q5 (Driver Speeding) shows a positive coefficient of 0.694, indicating an increased likelihood with behavior of drivers - speeding. Q6 (Driver Yielding) and Q8 (Presence of Raised Zebra Crossing) also exhibit positive coefficients, suggesting a higher likelihood with behavior of drivers – yielding and the presence of a raised zebra crossing, respectively. Notably, Q7 (Crossing Distance) and Q9 (Presence of Zebra Crossing) have negative coefficients, signifying a reduced likelihood with longer crossing distances and the presence of a zebra crossing. The 95% confidence intervals offer a range within which the true odds ratios are likely to fall. Overall, the results underline the importance of careful interpretation, considering both statistical significance and practical implications, and suggest avenues for further investigation and refinement of the predictive model.

Table 12 Table of Binary logistic regression model results.

Variables in the Equation

								95% C.I.for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1[a]	Q1	-.188	1.100	.029	1	.864	.828	.096	7.158
	Q2	-.223	.740	.091	1	.763	.800	.187	3.414
	Q3	-1.902	.981	3.760	1	.053	.149	.022	1.021
	Q4	-.611	1.008	.368	1	.544	.543	.075	3.914
	Q5	.694	1.070	.421	1	.516	2.002	.246	16.288
	Q6	1.215	.957	1.611	1	.204	3.370	.516	21.990
	Q7	-1.440	.875	2.713	1	.100	.237	.043	1.315
	Q8	1.250	1.256	.989	1	.320	3.489	.297	40.946
	Q9	-1.494	.754	3.931	1	.047	.224	.051	.983
	Constant	5.626	2.620	4.610	1	.032	277.669		

a. Variable(s) entered on step 1: Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9.

4.5 Integration of Survey and On-site Observation

The comprehensive integration of video recording analysis and survey data yields a holistic understanding of pedestrian decision-making in road-crossing scenarios. On-site observations highlight critical factors, such as the influence of vehicle types and the positive impact of larger pedestrian groups, as revealed by video analysis. Survey results affirm these findings, emphasizing the significance of vehicle type (QS1) and the presence of more pedestrians (QS2) as influential factors. The Exploratory Factor Analysis (EFA) further dissects latent constructs, with positive correlations indicating consistent responses. The logistic regression analysis links survey responses to decision outcomes, affirming the model's overall significance and goodness of fit through Omnibus Tests and Cox & Snell R Square and Nagelkerke R Square. The Classification Table underscores the model's accuracy, particularly in predicting road-crossing decisions ("CROSS"). Variable discussions highlight varying degrees of significance, emphasizing predictors like traffic volume, crossing distance, and the presence of zebra crossings, while acknowledging non-significance in other variables. This synthesis enriches the study by providing nuanced insights into the complex dynamics of pedestrian decision-making.

5. Conclusion

In conclusion, this report successfully achieved its objectives by identifying traffic-related factors influencing pedestrian crossing decisions and developing a pedestrian crossing decision model based on traffic conditions. The integration of on-site observations and survey data provided a comprehensive understanding of the multifaceted influences shaping pedestrian behaviors in urban settings. High traffic volume emerged as a discouraging factor, highlighting the need for strategies to reduce congestion or provide alternative pedestrian-friendly routes. Positive coefficients for vehicle speeding, yielding behavior, and raised crosswalks underscore their significant impact on increasing the likelihood of pedestrians crossing roads. The resulting pedestrian crossing decision model offers a robust foundation for evidence-based interventions, including traffic calming measures, awareness campaigns, and strategic placement of zebra crossings, to enhance pedestrian safety and traffic flow. In essence, this study not only met its objectives but also uncovered practical insights applicable to urban planning and traffic management. Ongoing research guided by these findings will contribute to creating safer and more pedestrian-friendly urban environments.

Acknowledgement

This research is made possible through monetary assistance by Universtii Tun Hussein Onn Malaysia under Tier 1 Research Grant Q346.

References

Amini, R. E., Katrakazas, C., & Antoniou, C. (2019). Negotiation and decision-making for a pedestrian roadway crossing: A literature review. In Sustainability (Switzerland) (Vol. 11, Issue 23). MDPI.

Bertulis, T., & Dulaski, D. M. (2014). Driver approach speed and its impact on driver yielding to pedestrian behavior at unsignalized crosswalks. In Transportation Research Record (Vol. 2464, pp. 46–51). National Research Council.

Camara, F., Romano, R., Markkula, G., Madigan, R., Merat, N., & Fox, C. (2018). Measuring Behavior.

Goddard, T., Kahn, K. B., & Adkins, A. (2015). Racial bias in driver yielding behavior at crosswalks. Transportation Research Part F: Traffic Psychology and Behaviour, 33, 1–6.

Hamed, M. M. (2001). Analysis of pedestrians' behavior at pedestrian crossings.

Malenje, J. O., Zhao, J., Li, P., & Han, Y. (2018). An extended car-following model with the consideration of the illegal pedestrian crossing. Physica A: Statistical Mechanics and Its Applications, 508, 650–661.

Masoud, M., Khaled, A., & Emhamed, M. A. (2016). Pedestrian Gap Acceptance and Crossing Decision outside Crossing Facilities along Urban Streets in Malaysia: A Case Study of Rughaya Street, Batu Pahat, Johor, Malaysia. The International Journal of Engineering And Information Technology (IJEIT), 3(1).

Nasrudin, N., Hashim, H., Khalid, N. S., Osman, N., & Hashim, S. F. (2020). Analysis of Pedestrian Crossing Behavior at Pedestrian Crossings in Shah Alam City, Malaysia. International Journal of Psychosocial Rehabilitation, 24(02), 1808–1815.

Obeid, H., Abkarian, H., Abou-Zeid, M., & Kaysi, I. (2017). Analyzing driver-pedestrian interaction in a mixed-street environment using a driving simulator. Accident Analysis and Prevention, 108, 56–65.

Papadimitriou, E., Lassarre, S., & Yannis, G. (2016). Pedestrian Risk Taking while Road Crossing: A Comparison of Observed and Declared Behaviour. Transportation Research Procedia, 14, 4354–4363.

Sucha, M., Dostal, D., & Risser, R. (2017). Pedestrian-driver communication and decision strategies at marked crossings. Accident Analysis and Prevention, 102, 41–50.

Sun, R., Zhuang, X., Wu, C., Zhao, G., & Zhang, K. (2015). The estimation of vehicle speed and stopping distance by pedestrians crossing streets in a naturalistic traffic environment. Transportation Research Part F: Traffic Psychology and Behaviour, 30, 97–106.

Tasic, I., Elvik, R., & Brewer, S. (2017). Exploring the safety in numbers effect for vulnerable road users on a macroscopic scale. Accident Analysis and Prevention, 109, 36–46.

Velasco, N., Farah, P. ;, Van Arem, H. ;, & Hagenzieker, B. ; (2017). Interactions between vulnerable road users and automated vehicles: A synthesis of literature and framework for future research. APA.

Wang, T., Jiang, X., Wang, W., & Cheng, Q. (2016). Analysis of Vehicle Speed Change at Non-signalized Crosswalks Based on Driving Behavior. Procedia Engineering, 137, 547–553.

Zhang, C., Zhou, B., Qiu, T. Z., & Liu, S. (2018). Pedestrian crossing behaviors at uncontrolled multi-lane mid-block crosswalks in developing world. Journal of Safety Research, 64, 145–154.