Accident Analysis and Prevention 129 (2019) 76–83

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Gap acceptance probability model for pedestrians at unsignalized mid-block crosswalks based on logistic regression

Jing Zhao*, Jairus Odawa Malenje, Yu Tang, Yin Han

University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai, PR China

A R T I C L E I N F O

Keywords:
Gap acceptance probability
Pedestrians
Mid-block crosswalks
Crossing behaviour
Logit model

A B S T R A C T

Gap acceptance represents a pedestrian's assessment of how safe it may be to use an available gap in traffic flow at a particular point in time. Though walking is a major component of urban mobility, the high rate of fatal interaction with motor vehicle traffic raises safety issues around how pedestrians decide to accept the available gap. This paper explored these interactions by modeling gap acceptance behavior at the midblock crosswalks. Unlike other pedestrian gap acceptance studies that focus on individual psychological and sociological factors that are difficult to control or manage, this study focused on six environmental factors that we considered important and as having the potential to affect the pedestrians' gap acceptance decision at the crosswalks, i.e. gap size, crossing distance, number of waiting pedestrians, waiting time, vehicle traffic volume and position of pedestrian (whether on street kerb or median). Video data was collected on pedestrian gap acceptance from 13 midblock crosswalk locations in Shanghai, China. A Logit model with 96% accuracy was developed to describe and predict the pedestrian gap acceptance behaviors. The results show that gap size and crossing distance have the highest effect on the pedestrian gap acceptance decision. Pedestrians waiting at the kerbside could confidently accept gaps (with a 95% probability) when the gap is longer than 2.2s, 5.9s, and 9.6s under the condition that the crossing distance is 4 m (one lane), 7.5 m (two lanes), and 11 m (three lanes), respectively while pedestrians waiting at the median could confidently accept gaps when the gap is longer than 1.6s, 5.3s, and 8.5s respectively under the same conditions. The recommendations on improving the crossing safety are proposed accordingly.

1. Introduction

Pedestrians are an integral part of the urban road transportation system. Unfortunately, as they share the roadway with motor vehicle traffic, there are increasing instances of tragic conflicts (Hatfield et al., 2007; Chandra et al., 2013; Jiang et al., 2015; Hochmuth and Houten, 2018; Shaon et al., 2018) some fatal. Especially, at midblock crosswalks, drivers consider pedestrian activity to be an obstacle to the uninterrupted flow of traffic (Zheng et al., 2016; Bertulis and Dulaski, 2018) while the pedestrians consider the flowing traffic an obstacle to their quest to reach the desired destination faster (Zhuang and Wu, 2011; Pawar et al., 2016). Unlike at the signalized crosswalks where the use of a road is enforced by the traffic signals (Lyons et al., 2001; Koh and Wong, 2014), the safety of pedestrians at unsignalized crosswalks largely depends on their capacity to correctly discern the traffic flow situation (Papadimitriou et al., 2012; Sun et al., 2015), choosing an appropriate time (Hatfield et al., 2007) crossing speed (Ishaque and Noland, 2008) and location to cross (Zhuang and Wu, 2012) and the motorists' propensity to yield when they should.

Pedestrian crosswalks have been widely implemented in many cities as a measure to address pedestrian safety against conflicts with vehicles (Yang et al., 2006; Kadali and Vedagiri, 2016). Although signalized crosswalks are more effective in dealing with the conflict between pedestrians and vehicles, the associated costs of setting up and running traffic signal control (Pawar and Patil, 2015) and the need for traffic flow efficiency (Zhuang and Wu, 2014; Pawar et al., 2016; Zhao et al., 2017) creates a constraint on the possible number of controlled crosswalks that can be set up on any roadway. By contrast, setting unsignalized crosswalks at mid-block is a more popular way.

As Petzoldt (2014) notes, risky crossing behaviors are difficult to modify, manage or control, this paper aims to establish a gap acceptance probability model for pedestrians at unsignalized mid-block crosswalk with the consideration of external factors. The objective of this paper is to assess pedestrian gap acceptance behavior considering

* Corresponding author.
E-mail addresses: jing_zhao_traffic@163.com (J. Zhao), jodawa@mmust.ac.ke (J.O. Malenje), ty1180517@hotmail.com (Y. Tang), hanyin2000@126.com (Y. Han).

https://doi.org/10.1016/j.aap.2019.05.012
Received 8 January 2019; Received in revised form 23 March 2019; Accepted 15 May 2019
Available online 22 May 2019
0001-4575/

Accident Analysis and Prevention 129 (2019) 76–83

external influencing factors. Therefore, there is a greater need to analyze pedestrian activity to ensure that the flow of human and vehicular traffic continues efficiently.

2. Literature review

Generally, the gap acceptance behavior of pedestrians can be explored broadly as both external factors and internal factors or personality traits. However, these factors have been discussed in most studies together in different combinations depending of the authors' focal point of interest. In the context of gap acceptance behavior, internal factors that affect pedestrian decision-making can be considered to revolve around individual psychological and sociological factors such as cautiousness, impatience, recklessness (Herrero-Fernández et al., 2016; Nor et al., 2017), different age groups (Oxley et al., 2005; Ferenchak, 2016; Lord et al., 2018; Shaaban et al., 2018), gender and age (Zhuang and Wu, 2012; Ferenchak, 2016; Lord et al., 2018; Venkata et al., 2018), temperament, assertiveness, and disability (Shaon et al., 2018; Zheng et al., 2016) among others. External factors include traffic density, waiting time and the available gap (Kadali and Vedagiri, 2013a; Iryo-Asano and Alhajyaseen, 2017a), traffic characteristics (platoon, gap size), vehicle characteristics (type, approach speed and distance), and road geometry characteristics (Hamidun et al., 2015; Iryo-Asano and Alhajyaseen, 2017a; Bertulis and Dulaski, 2018; Shaaban et al., 2018).

In traffic situations, pedestrians arrive at random at a crosswalk, then wait for an appropriate gap they deem safe to execute a crossing maneuver, dash across (Zhuang and Wu, 2011) or wait for a vehicle yield (Schroeder and Rouphail, 2011; Kadali and Vedagiri, 2013b; Papadimitriou et al., 2016a, b). In analytical studies on traffic flow done by Helbing et al. (2005), it was observed that when pedestrians chose a larger gap, vehicles only needed to slow down and do not necessarily have to stop. At times pedestrians fail to correctly observe traffic rules (Hatfield et al., 2007), misjudge the gap (Oxley et al., 2005; Schmidt and Färber, 2009) or simply hope that the driver will yield and end up attempting to cross when the available gap is unsafe. This can create a scary scenario described by Pawar et al. (2016) as a dilemma where they are not sure whether to proceed, stop or retreat because of the impending conflict with an oncoming vehicle. Traversing a multi-lane street is a bigger challenge for pedestrians, especially on roads with a fast flowing traffic that makes it harder to estimate the speed of oncoming vehicles on multiple lanes (Sun et al., 2015) and conflicts are potentially fatal if the crosswalk is unsignalized (Yang et al., 2006; Pawar and Patil, 2016; Shaaban et al., 2018). The risk is higher when a driver in one lane yields while another overtakes at high speed (Hochmuth and Houten, 2018). Some drivers also approach crosswalks aggressively while others ignore the pedestrian right of way thus endangering the lives of pedestrians who expect them to yield.

In their quest to describe the gap acceptance phenomena and predict the likelihood of pedestrians accepting available gaps or drivers yielding at unsignalized midblock crosswalks, researchers have developed models based on a mix of environmental and behavioral factors. Several studies (Schmidt and Färber, 2009; Schroeder and Rouphail, 2011; Kadali and Vedagiri, 2013a; Serag, 2014; Boroujerdian and Nemati, 2016; Shaaban et al., 2018) focused on behavioral-based predictive models for describing the gap acceptance in the interaction of drivers and pedestrian crossings. Some models consider behavioral characteristics, driver yielding behaviour, temporal and spatial gaps, and the challenges of searching for gaps in each lane in a multi-lane roadway affected their road crossing decisions (Pawar and Patil, 2015). Schmidt and Färber (2009) while considering group behaviour, age of pedestrians, gap sizes and street widths, developed a driver assistance system to protect pedestrians waiting at the kerbside by predicting their crossing intentions. According to (Harrell and Bereska, 1992; Oxley et al., 2005; Lobjois and Cavallo, 2007) older pedestrians had a challenge in correctly perceiving the available gap considering the speed of an oncoming vehicle. In their study conducted in Seattle, Washington state of the USA, Harrell and Bereska (1992) observed that older pedestrians and those with babies, were more cautious preferring lest risky longer gaps compared to other pedestrians. The accepted gaps with respect to conflicting traffic and pedestrians' changing crossing speed was also studied (Ishaque and Noland, 2008; Chandra et al., 2013; Iryo-Asano and Alhajyaseen, 2017b). Their study showed that accepted gaps decreased with conflicting traffic and crossing speed of the pedestrians and that the number of lanes on the road to be crossed affected the accepted gaps.

Other studies have developed gap acceptance models focused on pedestrian behavior at both signalized and unsignalized intersections (Sun et al., 2015; Bella et al., 2017; Iryo-Asano and Alhajyaseen, 2017a), and signalized (Zhuang et al., 2018). Lyons et al. (2001) developed an artificial neural network to model gap acceptance decisions for midblock signalized crosswalks with emphasis on pedestrian behaviour when presented with multiple crossing opportunities. In recent studies, Alver and Onelcin (2018) modelled pedestrian gap acceptance at overpasses using the logit model to estimate time gaps considering interactions of gender, age, and vehicle position, items carrying, and group size while Ramesh et al. (2018) considered the effect of land use on gap acceptance for pedestrian safety. However, these studies did not adequately consider other environmental factors and neither do they proffer interventions to make pedestrians safer. Such behavioral characteristics that inform the studies reviewed above and can affect the probability of gap acceptance are part of unpredictable human traits that vary from person to person at different points in time.

According to the above literature review, most of the existing studies focus on the internal factors. Whereas internal factors are important in pedestrian behavior, it is very difficult to change it. Therefore, in this study we present a different perspective of this phenomena by considering a combination of external factors (independent of pedestrian characteristics) that influence gap acceptance decisions at midblock unsignalized crosswalks. A gap acceptance probability model will be established to describe the influence of these factors on the crossing behavior of pedestrians at unsignalized mid-block crosswalks. An analysis of each factor is done to establish how it contributes to the pedestrian decision. This study goes beyond showing the influence of the identified factors on gap acceptance and presents scenarios that transportation administrators can consider to improve pedestrian safety.

3. Data collection

In this study, Gap acceptance (GA) is a binary outcome variable that indicates whether a pedestrian accepted the available time gap and crossed the roadway or rejected the gap and waited for the next available gap. It is the dependent variable in the study where 1 and 0 represents gaps accepted and gaps rejected respectively.

3.1. Analysis of potential influencing factors

We considered factors that affected gap acceptance and could be directly controlled and managed by transportation administrators. They were represented by the following six variables: gap size, crossing distance, platoon size, waiting time, traffic volume and position of pedestrian (whether on street kerb or median).

- Gap size, G_S (seconds) - It is an influencing scale-variable that represents the time the oncoming lead vehicle in a traffic stream will take to traverse the distance from point of sight at a particular instance in time to the crosswalk crossline. It can be computed as the ratio of distance over speed
- Crossing distance, D (meters) - It's the width of a roadway from nearside kerb (or median) to the far-side kerb depending on the road configuration, i.e. one way, or two-way. If the road has a median island, the crossing distance will be from kerb to median while if

there is no median, it is the distance from nearside kerb to the opposite kerb. It is a critical variable because pedestrians have to estimate how fast they can traverse the road-width before the next vehicle arrives.

- Waiting time, *T* (seconds) – this is the duration that a pedestrian takes to wait for a gap size that is safer to cross. It is measured from the time a pedestrian arrived at the kerb or median until he/she set foot on the street to cross. At times, if they take too long waiting, they may accept smaller gaps that are risky
- Volume of vehicles, V (veh/h) - the rate at which motor vehicles traverse a certain portion of a roadway. In this study it is the average hourly rate of vehicles passing the crosswalk spot within the duration of the survey each day. When the volume is high the available gaps tend to be smaller and unsafe. It is a discrete number
- Number of waiting pedestrians, *N* - It's the number of people waiting to cross the road at the kerb or median. It is the count at the point in time when there is an acceptable gap and before they begin to cross.
- Position of a pedestrian, *P* - Whether the pedestrian is waiting from the kerb or median, needs to be considered as an influencing variable. Those who wait at the median may feel increasingly unsafe if they have to wait too for the next gap available gap. There is a possibility that risky gaps can be accepted. It is a binary discrete variable where 1 and 0 represents a pedestrian waiting on a kerb and the median respectively.

3.2. The surveyed spots

A comprehensive survey was done by collecting data from 13 streets in Shanghai, China for 3 h each day over a period of 3 weekdays. These streets and the crosswalk spots were found to be ideal for the study because they cover the characteristics, geometric and management conditions of the potential influencing factors at mid-block crosswalks, as shown in Table 1. A total of 11,500 samples were collected. The percentage of females is 51.1%. The percentages of the teenagers (≤17 years old), the middle-age (18–59 years old) and the aged (≥60 years old) were about 13.5%, 70.6%, and 15.9%, respectively. Since the surveys were done during peak hours, the teenagers are identified by the school uniform and the schoolbag, while the aged are identified by their gait, white hair, the dress, and the stuff they carrying, such as the walking stick. Since the pedestrians can wait safely on the median island, they only need to wait the acceptable gap of one direction at once. It is just like the one-way street. Therefore, a two-way street with median island is treated as two one-way directions in this study, because it is very difficult to find a one-way street with four lanes in real world.

The situations considered were those where a car was approaching a crossing where a pedestrian was either waiting, approaching, or crossing the road. All the potential influencing factors mentioned in Section 3.1 were measured. To get a great view and ensure the accuracy of the survey, the equipment deployed for data collection included an Unmanned Aerial Vehicle (UAV), as shown in Fig. 1, and video data collection software, which has been successfully used in data collection in the authors' previous work (Zhao et al., 2016; Zhao and Liu, 2017; Zhao et al., 2018). The crossing distance was measured from the nearside (where the pedestrian is waiting) to the far-side (opposite side of the road) on a 1-way street or 1-lane 2-way street and nearside to Median Island or Median Island to far-side on a multi-lane 2-way street. If a pedestrian crossed and stopped at the median of a wide multi-lane street to wait for the next gap, the second crossing was considered a new crossing and recorded accordingly. Even those who arrived when the gap was acceptable and crossed were considered as waiting pedestrians but with 0 s waiting time.

Fig. 1. Example of the Unmanned Aerial Vehicle (UAV) video data collection.

4. Model development

A statistical analysis was performed on the data to establish trends in pedestrian gap acceptance behaviour and correlations between the variables. Pearson's correlation was used to establish the relationship between the variables and whether there is collinearity between them. Multiple Linear Regression (MLR) was used to evaluate the effect of the influencing variables on the outcome variable and develop a statistical model to explain goodness of fit in the model. A logit model was used to establish the probability of gap acceptance given the variables from different observations.

4.1. Introduction to the logistic model

The outcome variable in this study is binary where the pedestrian has to make one of two choices, whether to accept an available gap in traffic (1) and cross or reject the available gap (0) and wait for the next gap. For a binary response taking the values 1 and 0 (accepted and rejected respectively), the expected value is the probability that the

Table 1
Surveyed midblock crosswalks in Shanghai city.

Number of traffic lanes	Direction	Surveyed crosswalk	Existence and width of median island
1	One way	Nancang Rd. (Near Sinan Rd.)	No
2	One way	Yutian Rd. (Near Quyang Rd.)	No
		Yongxin Rd. (Gonghexin Rd.)	No
	Two way	Wulumuqi Rd. (Near Huashan Rd.)	No
		Linshi Rd. (South of the Nanhuayuan Rd.)	No
3	One way	Shangcheng Rd. (Near Dongfang Rd.)	No
	Two way	Linshi Rd. (North of the Nanhuayuan Rd.)	No
		Pingliang Rd. (Near Ninwu Rd.)	No
4	One way	Changshou Rd. (Near Wanhangdu Rd.)	Yes / 1.5 m
		Changshou Rd. (Near Jiaozhou Rd.)	Yes / 6.0 m
	Two way	Zhenhua Rd. (Near Xincun Rd.)	No
		Xincun Rd. (Near Zhenhua Rd.)	No
5	Two way	Yinkou Rd. (Jiamusi Rd.)	No

variable takes the value 1, i.e., the probability of accepting a gap. One of the assumptions of linear regression is that the relationship between predictor variables and the outcome variable is linear and for a linear regression model to be valid, the observed data should have a linear relationship in the form of $y = b + mx$ where, b is the constant (y intercept) and m is the coefficient of the predictor variable, x (or gradient).

In this study, the logistic regression model is the most suitable for modelling the binary outcome variable. The logistic model is derived from the Multiple linear regression function, U, written as the linear sum of β_0 plus the products of coefficients,? ?$_n$ and their corresponding X_n variables, the independent variables of interest, as shown in Eq. (1).

$$U = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \tag{1}$$

where n is the number of predictor variables,? ?$_0$ is a constant term representing the y intercept, and? ?$_n$ are unknown parameters associated with the predictor variables and quantifies their relationship with the outcome variable. Essentially then, U is an index that combines the predictor variables, X_n.

With the probability of accepting and rejecting gap denoted *p(U)* and *p(1-U)* respectively, then the probability of accepting a gap can be calculated using the Logit function, as shown in Eq. (2).

$$p(U) = \frac{1}{1 + e^{-U}} \tag{2}$$

Where $-\infty < U < \infty$, $0 < p < 1$ and e is the base of natural logarithms. As U approaches $-\infty$, p approaches 0 and as U approaches ∞, p approaches 1. Any *p(U)* value closer to 0 means that the probability of an event happening is very low while a value closer to 1 means that the event was very likely to occur.

The effect of selected variables on the pedestrian gap acceptance and road crossing behaviour is described with the help of multiple linear regression model. Model validation will be done using split sampling method by splitting the data randomly into two parts (75:25) - 75% for model development and 25% for validation data sets.

4.2. Tests of the logistic model application presuppositions

For multiple regression analysis to be employed, the data has to comply with the requirements for multi-collinearity between the predictor variables and linearity between outcome and predictor variables. In addition, there should be no outliers that can unduly influence the outcome variable.

4.2.1. Test for multi-collinearity between the predictor variables

Multi-collinearity exists when two or more of the predictors in an MLR model are moderately or highly correlated and has the effect of increasing the standard errors of the coefficients which in turn can affect the coefficients for some independent variables that may become significantly different from 0. When the standard errors are inflated, multi-collinearity can make some of the variables statistically insignificant when they could be significant. Table 2 is a summary of the MLR at 95% confidence levels. From the Table 2, the variance inflation factor (VIF) for all the independent variables was 1.00, which is an indicator that there was no significant linear correlation between them. Therefore, the data is suitable for logistic analysis.

Table 3
Box-Tidwell transformation.

	B	S.E.	Wald	df	Sig.	Exp(B)
G_S	3.145	.615	26.177	1	.000	23.230
D	−3.440	.858	16.088	1	.000	.032
T	.230	.124	3.450	1	.063	1.259
V	−.045	.061	.542	1	.462	.956
N	−.541	.484	1.246	1	.264	.582
P(1)	1.288	.165	60.580	1	.000	3.625
G_S * In_G_S	−.168	.224	.562	1	**.453**	.845
D * In_D	.191	.288	.440	1	**.507**	1.210
In_T * T	−.054	.039	1.864	1	**.172**	.948
In_V * V	.007	.009	.571	1	**.450**	1.007
In_N * N	.289	.258	1.255	1	**.263**	1.335
Constant	7.885	2.992	6.948	1	.008	2658.448

4.2.2. Test for linearity between outcome and predictor variables

A linear relationship between the predictor and outcome variables is one of the prerequisites for employing logistic regression model. The Box-Tidwell transformation was performed to test the linear relationship between individual continuous predictor variables and the outcome of the logit function. This is achieved by first converting the values of the predictor variables to natural logarithm, then computing the product of the two before a binary logit regression is done. Any transformed variable that is statistically significant at $p < 0.05$ fails the test of linear relationship with the outcome variable.

Therefore, from the results in Table 3, after the Box-Tidwell transformation all the transformed continuous predictor variables (gap size (ln_Gs, crossing distance, ln_D, waiting time, ln_T, volume of vehicles, ln_V and number of waiting pedestrians, ln_N) had a $p > 0.05$, indicating that they each had a linear relationship with the outcome variable.

Thus, having established that the data variables have no multi-collinearity and have individual linear relationship with the outcome variable, a logit model can be developed and used to predict gap acceptance accordingly.

4.3. Binary logistic model establishment

For developing the logit model, 75% of the data was used to establish the contribution of each of the predictor variables to the pedestrian gap acceptance decision. The rest 25% of the data collected was used to validate the model in Section 4.4. The descriptive statistics of the logit model are presented in Table 4.

From results in Table 4, four predictor variables (gap size, crossing distance, waiting time and pedestrian location) were found to be statistically significant but two variables (vehicle volume with a p-value of 0.339 and platoon size with a p-value of 0.957) failed the test of

Table 2
MLR coefficients for gap acceptance.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	S.E.	Beta			Tolerance	VIF
(Constant)	.745	.021		35.696	.000		
GS	.053	.001	.531	75.125	.000	.999	1.001
D	−.072	.001	−.541	−76.467	.000	1.000	1.000
T	.001	.001	.014	1.994	.046	1.000	1.000
V	1.912E-7	.000	.000	.003	.997	1.000	1.000
N	.002	.003	.004	.507	.612	1.000	1.000
P	.036	.007	.036	5.120	.000	1.000	1.000

Table 4
Variables in the logit model.

	B	S.E.	Wald	Df	Sig.	Odds ratio
Constant	6.365	.494	165.741	1	.000	581.107
GS	2.678	.102	692.611	1	.000	14.550
D	−2.846	.107	705.033	1	.000	.058
T	.058	.012	21.677	1	.000	1.059
V	.001	.001	.915	1	.339	1.001
N	.004	.066	.003	1	.957	1.004
P	−1.273	.158	64.686	1	.000	.280

statistical significance. A plausible explanation for this turn or results could be due to the location of the crosswalk at the midblock where there is a constant flow of vehicle traffic. It could also be due to the effects of traffic flows at the intersections on both sides of the roadway arising from signal timings.

The odds ratio is a statistic that represents the change in odds of a predictor variable influencing the change in the outcome variable and its reference pivot point is 1. Given that the value of 1 is the threshold at which the direction of the effect changes, then the odds ratios of 1 (or near 1) indicates the least effect (or none) on the outcome variable by the predictor. In this study, we can say that the odds of a pedestrian accepting a gap are 14.55 times higher with each increase in the available gap size and 0.058 lower with each unit increase in crossing distance because it's less than 1. However, there will be minimal effect with pedestrian waiting time, vehicle volume and pedestrian platoon sizes because their odds rations are very close to 1.

Based on the results presented above, and omitting the variables that were found to be statistically insignificant, the regression model can be written as Eq. (3). The probability of pedestrians in this study accepting the available gaps can be expressed through the probability function, as shown in Eq. (4).

$$U = 6.365 + (2.678 * G_s) - (2.846 * D) + (.058 * T) - (1.273 * P) \quad (3)$$

$$p(G_A) = \frac{1}{1 + e^{-U}} \quad (4)$$

4.4. Validity of the model

The validation of this model was done using the rest 25% of the data collected. As shown in Table 5, there are four possible situations, namely, the result of the proposed model is the same as the surveyed result (both accept or reject the gap) and the result of the proposed model is different from the surveyed result (accept in the model but reject in reality, or reject in the model but accept in reality). In the comparative analysis, the model prediction for accepted gaps was based on a probability cut-off of 0.5 which implies that any probability value below 0.5 was considered a gap rejected. The results show that 2783 times out of 2876 observations, the model concurred with the observed gap acceptance behaviour (96.77% accuracy). Paired *t*-test results (see Table 6) further show no significant difference between results estimated by the proposed model and that from simulation (p-value = 0.604 > 0.05), indicating that the accuracy of the proposed gap acceptance model is acceptable.

Table 5
Comparative analysis of gap acceptance model against observations.

	Observed	Proposed model		
		True	Wrong	Accuracy
Gap accepted	1895	1851	44	97.68%
Gap Rejected	981	932	49	95.01%
Total	2876	2783	93	96.77%

5. Discussion

Sensitivity analysis is used in statistical modeling to analyze how the different values of a set of predictor variables affect the outcome variable under certain specific conditions. This study examined the influence of six variables on pedestrian gap acceptance at the midblock unsignalized crosswalks in Shanghai city. Two of the variables were found not to be statistically significant and were not used to develop the logit model. From the four variables that were statistically significant, two variables (pedestrian waiting time and pedestrian position) were weakly correlated with the outcome variable while gap size and crossing distance were strongly correlated with the outcome variable. This analysis examines how the variables affect the outcome variable. The basic input parameters are: gap size (8 s), crossing distance (10 m), waiting time (60 s) and pedestrian position (1, kerb).

5.1. Influence of gap size and crossing distance

Gap size and crossing distance are important and inextricably connected influence factors on whether or not an available gap will or can be accepted. While mentally considering the gap size of the oncoming vehicle, the pedestrian has to also think about how long it will take to start and finish crossing. Fig. 2 shows the general relationship between the gap acceptance probability and the combination of gap size and crossing distance. Fig. 3 further shows the gap acceptance probability against crossing distance under a range of fixed gap size cases.

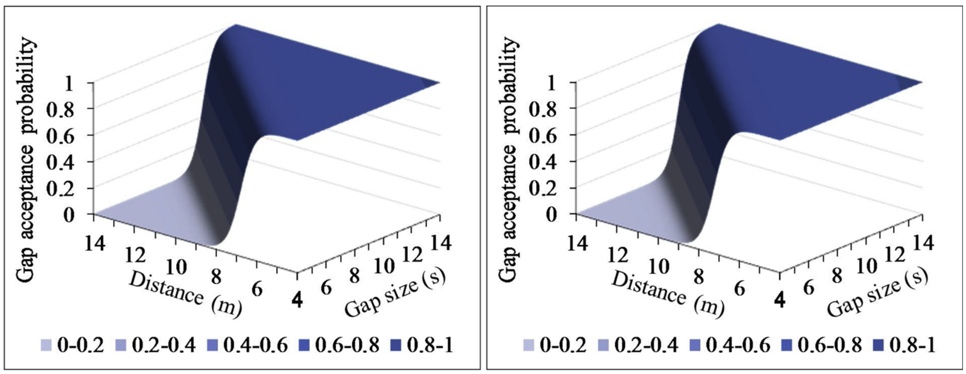
The Fig. 2 shows that longer crossing distances require an equally larger gap size for the pedestrian to accept the available gap. For instance, using 0.5 as the predicted probability threshold for gap acceptance (Petzoldt, 2014), a distance of 5 m required a gap of 2.1 s, a 7 m and 10 m distance, requires 4.3 s and 7.4 s respectively for pedestrians waiting by the kerbside (Fig.2a) while the same distance needed gaps of 1.6 s, 3.8 s and 6.9 s respectively for those waiting at the median (Fig. 2b). Fig. 3 gives a different perspective, showing the critical distances where given gap sizes will be acceptable to pedestrians when the probability is 50%. The probability of accepting a gap size of 4 s, 6 s or 7 s is high (50% or more) if the crossing distance is less 6.8 m, 8.6 m and 9.6 m respectively at the kerbside (Fig. 3a) while for pedestrians waiting at the median, the same gap sizes can be accepted if the distances are 7.2 m, 9.1 m and 10 m respectively (Fig. 3b). It can be seen that under the same condition, the gap acceptance probability of pedestrians waiting at the median is slightly higher than that waiting at the kerbside (3% on average).

Accepting the small gap for the shorter distance (most likely on a one-way single-lane street) could be attributed to the normally slow speed of oncoming vehicles on a single-lane street as a precaution against the ever-present risk of pedestrians emerging from behind parked cars. For the larger crossing distances, pedestrians need bigger gaps because ordinarily, traffic flow is faster on a multi-lane roads (Bertulis and Dulaski, 2018), as impatient drivers eager to reach their destinations tend to yield less to pedestrians and overtake the slow moving vehicles. Pedestrians waiting at the kerbside could confidently accept gaps (with a 95% probability) when the gap is longer than 2.2 s, 5.9 s, and 9.6 s under the condition that the crossing distance is 4 m (one lane), 7.5 m (two lanes), and 11 m (three lanes), respectively while pedestrians waiting at the median could confidently accept gaps when the gap is longer than 1.6 s, 5.3 s, and 8.5 s respectively under the same conditions.

For different crossing distances there is a turning point for each gap size where the probability of gap acceptance dips from a high of 95% to less than 5%. This is an indicator that beyond that point, even though these are designated crosswalks, pedestrians found it inherently risky to attempt to cross and traverse that distance unless some calming interventions are implemented. Possible interventions for streets wider than 12 m could include median islands, so that pedestrians can traverse the crosswalk in two stages. Otherwise, a signal control or a pedestrian

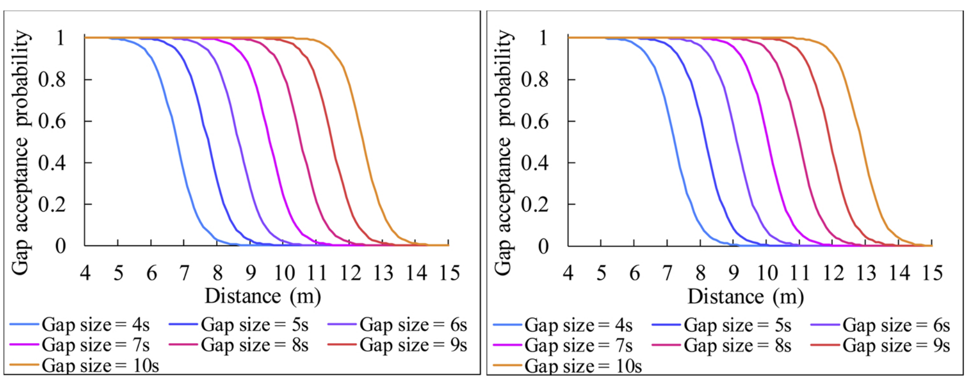
Table 6
Paired *t*-test results.

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Observed - Proposed model	−.00174	.17985	.00335	−.00831	.00484	−.518	2875	.604



(a) Kerbside waiting (b) Median waiting

Fig. 2. Effect of gap size and crossing distance on gap acceptance.



(a) Kerbside waiting (b) Median waiting

Fig. 3. Gap acceptance probability against crossing distance under different gap size cases.

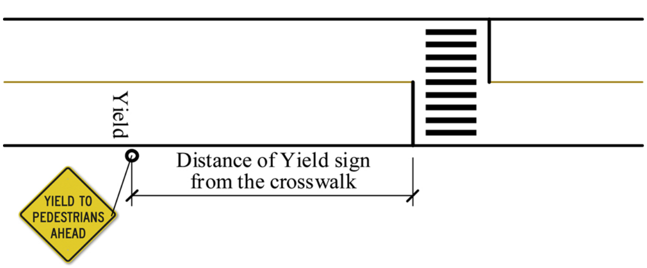


Fig. 4. Measures to remind the driver to yield to the pedestrian based on the gap size analysis.

Table 7
Location of Yield sign from crosswalk (m).

		Road Design speed				
		60km/h	50km/h	40km/h	30km/h	20km/h
Crossing distance	4m	48	40	32	24	16
	7.5 m	270	225	180	135	90
	11 m	480	400	320	240	160

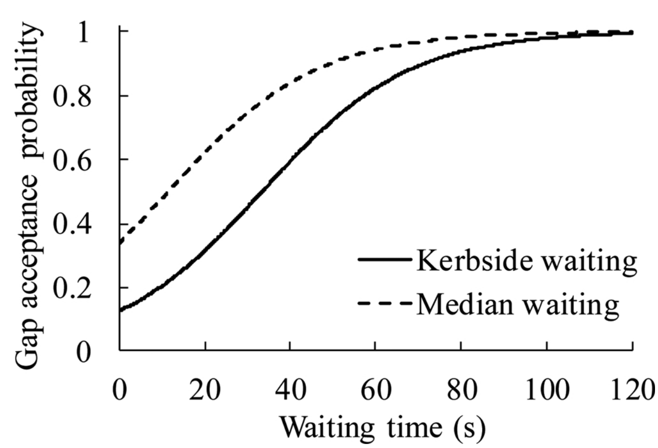


Fig. 5. Pedestrian waiting time.

overpass is required to reduce interruption on traffic flow.

Moreover, given that this is a designated crosswalk where pedestrians have right of way, to improve their crossing safety, a Yield reminder (including a ground-mounted sign and a word pavement marking), can be installed at a distance equivalent to a time to arrival that will be applicable to each street (the point where probability of gap acceptance is 30%) from the crosswalk considering the road segment's design speed (the expected speed of the approaching vehicle in compliance with the applicable legal provisions) and the road width (pedestrian's crossing distance). The function of the Yield reminder is to remind the driver that pedestrians may cross under the condition, and they should yield to them to avoid the potential conflict. The recommended ground-mounted yield reminder style and the distance between the yield reminder and the crosswalk are shown in Fig. 4 and Table 7, respectively.

5.2. Influence of waiting time and pedestrian position

Pedestrians arrive and wait for the next available gap if there is vehicle traffic flow. A pedestrian waits for the next available gap on the nearside kerb or median if the road has multiple lanes. The position of a pedestrian can also have a bearing on the crossing distance and the time it will take to cross the street when a gap becomes available. If the flow is fast, consistent and heavy, they may wait longer but if it is intermittent and slow, gaps of various sizes will be many and they can take advantage and cross.

Fig. 5 shows the effect of waiting times on the probability of gap acceptance. For pedestrians waiting at the kerbside, they are more likely (over 50%) to accept available gaps after waiting for 33 s and above but at the median, they are likely (over 50%) to wait for no more than 12 s. Although they have right of way at the crosswalk, pedestrians are still careful about the risk of crossing by waiting longer, especially at the kerbside. Pedestrians waiting at the median are more likely to accept available gaps than those waiting at the kerbside. For the pedestrians waiting at the median, they may feel unsafety. Therefore, they prefer accepting a smaller gap than waiting longer.

Moreover, with the increase of waiting time, the gap acceptance probability increases under the same condition. It indicates that pedestrians may feel impatient with the increase of waiting time. It's dangerous because drivers don't know these pedestrians lose patient. Provision of a pedestrian-activated button signal control can help alert drivers to yield especially for those waiting at the median.

6. Conclusion

The main objective of this study was to establish the factors that influence the pedestrian gap acceptance and road crossing behavior at the midblock crosswalks. Unlike other pedestrian gap acceptance studies that focus on a mix of pedestrian behaviour that cannot be controlled directly and environmental factors, this study focused on six environmental factors that we considered important and as having the potential to affect the pedestrian gap acceptance decision of pedestrians at the crosswalks: temporal gap size, crossing distance, platoon size, waiting time, traffic volume and position of pedestrian (whether on street kerb or median). Observation data was collected from 13 streets ranging from single lane one-way to 5 lane bi-direction roadway. Multiple Linear Regression and Logit models were used to analyze the data. The model developed had a 96% accuracy in predicting the gap acceptance behavior. The sensitivity analysis was conducted and the following observations can be made:

(1) A logistic regression model was established to estimate the gap acceptance probability of pedestrians at unsignalized mid-block crosswalks. Four external factors, including gap size, crossing distance, waiting time and pedestrian location, were found to be statistically significant in determining whether pedestrians will either accept or reject an available gap.
(2) Gap size and crossing distance had the highest effect on the pedestrian gap acceptance decision. Pedestrians waiting at the kerbside could confidently accept gaps (with a 95% probability) when the gap is longer than 2.2 s, 5.9 s, and 9.6 s under the condition that the crossing distance is 4 m (one lane), 7.5 m (two lanes), and 11 m (three lanes), respectively while pedestrians waiting at the median could confidently accept gaps when the gap is longer than 1.6 s, 5.3 s, and 8.5 s respectively under the same conditions.
(3) To improve the crossing safety, the installation of "YIELD" signs at designated locations away from the stop-line considering the design speed and width of the road segment is recommended to remind the driver that pedestrians may cross under the condition. Moreover, crossing distance of more than 12 m was found to be risky and required appropriate interventions such median island.

The findings tie up with what happens in real life traffic where pedestrians use the gap size and crossing distance to assess the viability of an available gap. Each pedestrian is inherently aware of their ability to successfully use the gap available considering the time it will take to safely traverse the distance before the oncoming vehicle arrives. This study can be extended to consider the pedestrian gap acceptance for illegal pedestrian crosswalks.

Acknowledgement

The research is supported by the National Natural Science Foundation of China under grant No. 51608324.

References

Alver, Y., Onelcin, P., 2018. Gap acceptance of pedestrians at overpass locations. Transp. Res. Part F Traffic Psychol. Behav. 56, 436–443.
Bella, F., Natale, V., Silvestri, M., 2017. Driver-pedestrian interaction under different road environments. Transp. Res. Procedia 27 (2017), 148155.
Bertulis, T., Dulaski, D.M., 2018. Driver approach speed and its impact on driver yielding to pedestrian behavior at unsignalized crosswalks. Transp. Res. Rec. 2464 (1), 46–51.
Boroujerdian, A.M., Nemati, M., 2016. Pedestrian gap acceptance logit model in unsignalized crosswalks conflict zone. Int. J. Transp. Eng. 4 (2), 87–96.
Chandra, S., Rastogi, R., Das, V.R., 2013. Descriptive and parametric analysis of pedestrian gap acceptance in mixed traffic conditions. Ksce J. Civ. Eng. 18 (1), 284–293.
Ferenchak, N.N., 2016. Pedestrian age and gender in relation to crossing behavior at midblock crossings in india. J. Traffic Transp. Eng. (English Ed.) 3 (4), 345–351.
Hamidun, R., Kordi, N.E., Endut, I.R., Ishak, S.Z., Faudzi, M., Yusof, M., 2015. Estimation of illegal crossing accident risk using stochastic petri nets. J. Eng. Sci. Technol. 10, 81–93.
Harrell, W.A., Bereska, T., 1992. Gap acceptance by pedestrians. Percept. Mot. Skills 75 (2), 432–434.
Hatfield, J., Fernandes, R., Job, R.S., Smith, K., 2007. Misunderstanding of right-of-way rules at various pedestrian crossing types: observational study and survey. Accid. Anal. Prev. 39 (4), 833–842.
Helbing, D., Jiang, R., Treiber, M., 2005. Analytical investigation of oscillations in intersecting flows of pedestrian and vehicle traffic. Phys. Rev. E 72 (4), 046130.
Herrero-Fernández, D., Macía-Guerrero, P., Silvano-Chaparro, L., Merino, L., Jenchura, E.C., 2016. Risky behavior in young adult pedestrians: personality determinants, correlates with risk perception, and gender differences. Transp. Res. Part F 36, 14–24.
Hochmuth, J., Houten, R.V., 2018. Influence of advanced placement of the in-street sign gateway on distance of yielding from the crosswalk. Transp. Res. Rec. 2672 (35), 13–20.
Iryo-Asano, M., Alhajyaseen, W., 2017a. Consideration of a pedestrian speed change model in the pedestrian–vehicle safety assessment of signalized crosswalks. Transp. Res. Procedia 21, 87–97.
Iryo-Asano, M., Alhajyaseen, W., 2017b. Modeling pedestrian crossing speed profiles considering speed change behavior for the safety assessment of signalized intersections. Accid. Anal. Prev. 108, 332–342.
Ishaque, M.M., Noland, R.B., 2008. Behavioural issues in pedestrian speed choice and street crossing behaviour: a review. Transp. Rev. 28 (1), 61–85.
Jiang, X., Wang, W., Bengler, K., Guo, W., 2015. Analyses of pedestrian behavior on midblock unsignalized crosswalk comparing Chinese and German cases. Adv. Mech. Eng. 7 (11), 168781401561046.
Kadali, B.R., Vedagiri, P., 2013a. Effect of vehicular lanes on pedestrian gap acceptance behaviour. Procedia - Soc. Behav. Sci. 104, 678–687.
Kadali, B.R., Vedagiri, P., 2013b. Modelling pedestrian road crossing behaviour under mixed traffic condition. Eur. Transp.-Trasporti Europei 55, 1–17.
Kadali, B.R., Vedagiri, P., 2016. Pedestrian crossing treatment warrants for midblock crosswalks under mixed traffic conditions. Transp. Res. Rec. 2581, 145–153.
Koh, P.P., Wong, Y.D., 2014. Gap acceptance of violators at signalised pedestrian crossings. Accid. Anal. Prev. 62, 178–185.
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. Accid. Anal. Prev. 39 (5), 934–943.
Lord, S., Cloutier, M.-S., Garnier, B., Christoforou, Z., 2018. Crossing road intersections in old age—with or without risks? Perceptions of risk and crossing behaviours among the elderly. Transp. Res. Part F Traffic Psychol. Behav. 55, 282–296.
Lyons, G., Hunt, J., Mcleod, F., 2001. A neural network model for enhanced operation of midblock signalled pedestrian crossings. Eur. J. Oper. Res. 129 (2), 346–354.
Nor, S.N.M., Daniel, B.D., Hamidun, R., Bargi, Wa.A., Rohani, M.M., Prasetijo, J., Aman, M.Y., Ambak, K., 2017. Analysis of pedestrian gap acceptance and crossing decision in kuala lumpur. MATEC Web of Conferences 103, 08014.
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. Accid. Anal. Prev. 37 (5), 962–971.
Papadimitriou, E., Lassarre, S., Yannis, G., 2016a. Introducing human factors in pedestrian crossing behaviour. Transp. Res. Part F 36, 69–82.
Papadimitriou, E., Lassarre, S., Yannis, G., Tselentis, D.I., 2016b. Road, traffic, and human factors of pedestrian crossing behavior: integrated choice and latent variables models. Transp. Res. Rec. 2586 (1), 28–38.
Papadimitriou, E., Yannis, G., Golias, J., 2012. Analysis of pedestrian exposure to risk in relation to crossing behavior. Transp. Res. Rec. 2299 (1), 79–90.
Pawar, D.S., Kumar, V., Singh, N., Patil, G.R., 2016. Analysis of dilemma zone for pedestrians at high-speed uncontrolled midblock crossing. Transp. Res. Part C Emerg. Technol. 70, 42–52.
Pawar, D.S., Patil, G.R., 2015. Pedestrian temporal and spatial gap acceptance at midblock street crossing in developing world. J. Saf. Res. 52, 39–46.
Pawar, D.S., Patil, G.R., 2016. Critical gap estimation for pedestrians at uncontrolled midblock crossings on high-speed arterials. Saf. Sci. 86, 295–303.
Petzoldt, T., 2014. On the relationship between pedestrian gap acceptance and time to arrival estimates. Accid. Anal. Prev. 72, 127–133.
Ramesh, A., Ashritha, K., Kumar, M., 2018. Development of model for pedestrian gap based on land use pattern at midblock location and estimation of delay at intersections. J. Inst. Eng. (India): Ser. A 99 (3), 413–422.
Schmidt, S., Färber, B., 2009. Pedestrians at the kerb – recognising the action intentions of humans. Transp. Res. Part F Traffic Psychol. Behav. 12 (4), 300–310.
Schroeder, B.J., Rouphail, N.M., 2011. Event-based modeling of driver yielding behavior at unsignalized crosswalks. J. Transp. Eng. 137 (7), 455–465.
Serag, M.S., 2014. Modelling pedestrian road crossing at uncontrolled midblock locations in developing countries. Int. J. Civ. Struct. Eng. 4 (3), 274–285.
Shaaban, K., Muley, D., Mohammed, A., 2018. Analysis of illegal pedestrian crossing behavior on a major divided arterial road. Transp. Res. Part F Traffic Psychol. Behav. 54, 124–137.
Shaon, M.R.R., Schneider, R.J., Qin, X., He, Z., Sanatizadeh, A., Flanagan, M.D., 2018. Exploration of pedestrian assertiveness and its association with driver yielding behavior at uncontrolled crosswalks. Transp. Res. Rec. 2672 (35), 69–78.
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. Transp. Res. Part F Traffic Psychol. Behav. 30, 97–106.
Venkata, K., Ravishankar, R., Nair, P.M., 2018. Pedestrian risk analysis at uncontrolled midblock and unsignalised intersections. J. Traffic Transp. Eng. (English Ed.) 5 (2), 137–147.
Yang, J., Deng, W., Wang, J., Li, Q., Wang, Z., 2006. Modeling pedestrians' road crossing behavior in traffic system micro-simulation in china. Transp. Res. Part A Policy Pract. 40 (3), 280–290.
Zhao, J., Li, P., Zheng, Z., Han, Y., 2018. Analysis of saturation flow rate at tandem intersections using field data. Iet Intell. Transp. Syst. 12 (5), 394–403.
Zhao, J., Liu, Y., 2017. Safety evaluation of intersections with dynamic use of exit-lanes for left-turn using field data. Accid. Anal. Prev. 102, 31–40.
Zhao, J., Liu, Y., Wang, T., 2016. Increasing signalized intersection capacity with unconventional use of special width approach lanes. Comput. Civ. Infrastruct. Eng. 31 (10), 794–810.
Zhao, J., Ma, W.J., Li, P., 2017. Optimal design of midblock crosswalk to achieve trade-off between vehicles and pedestrians. J. Transp. Eng. Part A-Syst. 143 (1), 04016003.
Zheng, Y., Elefteriadou, L., Chase, T., Schroeder, B., Sisiopiku, V., 2016. Pedestrian traffic operations in urban networks. Transp. Res. Procedia 15, 137–149.
Zhuang, X., Wu, C., 2011. Pedestrians' crossing behaviors and safety at unmarked roadway in china. Accid. Anal. Prev. 43 (6), 1927–1936.
Zhuang, X., Wu, C., 2012. The safety margin and perceived safety of pedestrians at unmarked roadway. Transp. Res. Part F Traffic Psychol. Behav. 15 (2), 119–131.
Zhuang, X., Wu, C., 2014. Pedestrian gestures increase driver yielding at uncontrolled mid-block road crossings. Accid. Anal. Prev. 70, 235–244.
Zhuang, X., Wu, C., Ma, S., 2018. Cross or wait? Pedestrian decision making during clearance phase at signalized intersections. Accid. Anal. Prev. 111, 115–124.